

GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.

## OM Protein - protein search, using sw model

Run on: April 8, 2004, 13:14:58 ; Search time 56 Seconds  
 Perfect score: 754 (without alignment)  
 Sequence: 1 MAMVSAWSWLYLWISACAM.....SGNMCATGNGKIKTRIHPRT 133

Title: US-09-763-335-2  
 Scoring table: BLOSUM62  
 Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: \_geneseqP1980s:\*

2: \_geneseqP1990s:\*

3: \_geneseqP2000s:\*

4: \_geneseqP2001s:\*

5: \_geneseqP2002s:\*

6: \_geneseqP2003as:\*

7: \_geneseqP2003bs:\*

8: \_geneseqP2004s:\*

RESULT 1

ID AAV70400

standard; protein; 133 AA.

XX

AAV70400;

AC

21-JUN-2000 (first entry)

DT

Human cell-signalling protein-2.

XX

DE

Human cell-signalling protein-2;

neoplastic disorder;

smooth muscle disorder;

vesicle trafficking disorder;

asthma; emphysema;

rheumatoid arthritis;

HIV; human immunodeficiency virus;

osteoporosis; multiple sclerosis;

arthroscclerosis; diabetes mellitus;

psoriasis; viral infection;

irritable bowel syndrome;

bacterial infection; fungal infection;

cytostatic; antiarthritic; antiasthmatic;

immunosuppressive;

antiarterioclerotic; anti-HIV; antidiabetic;

antipsoriatic; antimicrobial; drug screening.

XX

OS Homo sapiens.

XX

Key Peptide

1 .. 25

/label= Signal\_Peptide

64

/note= "Protein kinase C phosphorylation site"

FT

Modified-site

77

/note= "Protein kinase C phosphorylation site"

FT

Modified-site

83

/note= "Casein kinase II phosphorylation site"

FT

Modified-site

109

/note= "Casein kinase II phosphorylation site"

FT

Modified-site

126

/note= "Protein kinase C phosphorylation site"

XX

WO200011169-A2.

## ALIGNMENTS

26	87	11.5	188	5	ABP67052
27	86.5	11.5	2703	4	ABB60266
	28	86	11.4	4	Aau16948
	29	86	11.4	500	Aaw94496
	30	86	11.4	503	AAY68936
	31	86	11.4	504	AAY68935
	32	86	11.4	505	AAY68934
	33	86	11.4	506	AAY68933
	34	86	11.4	507	AAY68932
	35	86	11.4	508	AAY68931
	36	86	11.4	524	AAY68942
	37	86	11.4	525	AAY68941
	38	86	11.4	526	AAY68940
	39	86	11.4	527	AAY68939
	40	86	11.4	528	AAY68938
	41	86	11.4	528	ABR39450
	42	86	11.4	529	Aay68937
	43	86	11.4	659	Aaw44497
	44	86	11.4	677	Aay68948
	45	86	11.4	678	Aay68947

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## \* Query

No.

Score

Match

Length

DB

ID

Description

Result No.	Score	Query	Match	Length	DB	ID	Description	Location/Qualifiers
1	754	100.0	133	3	AAV70400		Aay70400 Human cell	XX
2	754	100.0	133	4	AAW76667		Human Pro	XX
3	754	100.0	190	4	AAW76651		Human Pro	XX
4	746	98.9	133	4	ABU22838		Human kid	FT
5	432	57.3	131	4	AAW73919		Human Pro	FT
6	432	57.3	242	4	AAW40977		Human Pro	FT
7	241	32.0	121	3	AAW44958		Human sec	FT
8	241	32.0	132	3	AAW53016		Human sec	FT
9	240.5	31.9	89	4	ABU22839		Abu52339 Human kid	FT
10	221.5	29.4	115	4	AAW24360		Aam24360 Human EST	FT
11	150	19.9	29	2	AAW13167		Aay13167 Human sec	FT
12	110.5	14.7	1202	7	ADB60833		Ad60833 Rat Prote	FT
13	96	12.7	1964	2	AAW95557		Aaw95557 Mus muscu	FT
14	95.5	12.7	2331	7	ADB63713		Ade63713 Rat Prote	FT
15	95.5	12.7	2531	7	ADB63705		Ade63705 Rat Prote	FT
16	95.5	12.7	2531	7	ADB63709		Ade63709 Rat Prote	FT
17	92.5	12.7	2331	7	ADB63701		Ade63701 Rat Prote	XX
18	92.5	12.3	2146	4	ABB62317		Abb62317 Drosophil	FT
19	91.5	12.1	233	2	AAV29053		Aay29053 T. gondii	XX
20	91.5	12.1	233	2	AAU25524		Aau25524 T. gondii	XX
21	87	11.5	188	4	AAU18112		Aau18112 Novel hum	PR
22	87	11.5	188	4	AAU17021		Aau17021 Human nov	PR
23	87	11.5	188	4	ABB10465		Abb10465 Human CDN	XX
24	87	11.5	188	4	AAU19912		Aau19912 Novel hum	XX
25	87	11.5	188	5	ABJ05739		Abj05739 Novel hum	PI

Tang YT, Corley NC, Patterson C, Guegler KJ, Baughn MR,

Baughn MR;

XX WPI; 2000-246562/21.  
DR N PSDB; AA51528.

XX New human cell signaling proteins and polynucleotides useful for diagnosis, prevention and trafficking of neoplastic, neurological, immunological, vesicle trafficking and smooth muscle disorders.

XX Claim 1; Page 60; 62pp; English.

CC The present sequence is human cell-signalling protein-2 (CSIG-2) involved in regulation of cell proliferation, differentiation and gene transcription. CSIG can be used in the diagnosis and treatment of diseases associated with expression of CSIG. These diseases include neoplastic, neurological, immunological, vesicle trafficking and smooth muscle disorders, including HIV, rheumatoid arthritis, asthma, arterioclerosis, diabetes mellitus, emphysema, irritable bowel syndrome, multiple sclerosis, osteoporosis, psoriasis and infections including viral, bacterial and fungal. CSIG and its catalytic or immunogenic fragments are useful for drug screening using libraries of compounds

SQ Sequence 133 AA;  
Query Match 100.0%; Score 754; DB 3; Length 133;  
Best Local Similarity 100.0%; Pred. No. 1.9e-69;  
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAMYSAMSWLYLWISACMLLCHGSLIGHTFQQLHLRPEGGTCEVIAHRCNNKRILE 60  
Db 1 MAMYSAMSWLYLWISACMLLCHGSQHTFQQLHLRPEGGTCEVIAHRCNNKRILE 60  
Qy 61 RSQTIVKCSCLPGKVAGTPRNPSCYDASIVIGKWCENMPCLGECKTLPDNSGMICAT 120  
Db 61 RSQTIVKCSCLPGKVAGTPRNPSCYDASIVIGKWCENMPCLGECKTLPDNSGMICAT 120

Qy 121 GNKIKTRIHPRT 133  
Db 121 GNKIKTRIHPRT 133

RESULT 2  
AM78667  
ID AM78667 standard; protein; 133 AA.

AC AAM78667;

XX DT 06-NOV-2001 (first entry)

XX DE Human protein SEQ ID NO 1329.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematoopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PP 05-FEB-2001; 20001WO-US004098.

XX PR 03-FEB-2000; 2000US-00496914.  
PR 27-APR-2000; 2000US-00560875.  
PR 20-JUN-2000; 2000US-00598075.  
PR 19-JUL-2000; 2000US-00620325.  
PR 01-SEP-2000; 2000US-00654936.  
PR 15-SEP-2000; 2000US-00663551.  
PR 20-OCT-2000; 2000US-00693325.  
PR 30-NOV-2000; 2000US-00728422.

XX PA (HYSEQ INC.

XX PI Tang YT, Liu C, Dumanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Weijhrman T, Goodrich R;  
XX WPI; 2001-476283/51.  
DR N-PSDB; AA51800.

XX PS Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.

XX XX PS Claim 20; Page 3574; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM878323-AAM8302) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g., stem cell growth factor activity, haematoopoiesis regulating activity, tissue growth factor activity and/or immunomodulatory activity and/or activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAR52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

XX XX SQ Sequence 133 AA;  
Query Match 100.0%; Score 754; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 1.9e-69;  
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX XX Qy 1 MAMVSANSWLYLWISACMLLCHGSQHTFQQLHLRPEGGTCEVIAHRCNNKRILE 60  
Db 1 MAMVSANSWLYLWISACMLLCHGSQHTFQQLHLRPEGGTCEVIAHRCNNKRILE 60  
Qy 61 RSOTIVKCSCLPGKVAGTPRNPSCYDASIVIGKWCENMPCLGECKTLPDNSGMICAT 120  
Db 61 RSOTIVKCSCLPGKVAGTPRNPSCYDASIVIGKWCENMPCLGECKTLPDNSGMICAT 120  
Qy 61 RSOTIVKCSCLPGKVAGTPRNPSCYDASIVIGKWCENMPCLGECKTLPDNSGMICAT 120  
Db 61 RSOTIVKCSCLPGKVAGTPRNPSCYDASIVIGKWCENMPCLGECKTLPDNSGMICAT 120  
Qy 121 GNKIKTRIHPRT 133  
Db 121 GNKIKTRIHPRT 133

RESULT 3  
ARM79651  
ID AAM79651 standard; protein; 190 AA.

XX AC AAM79651;

XX DT 06-NOV-2001 (first entry)

XX DE Human protein SEQ ID NO 3297.  
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematoopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX OS Homo sapiens.

XX PN WO200157190-A2.  
XX PD 09-AUG-2001.  
XX PP 05-FEB-2001; 20001WO-US004098.  
XX PR 03-FEB-2000; 2000US-00496914.  
XX PD 09-AUG-2001.  
XX PP 05-FEB-2001; 2000US-0046914.  
XX PR 27-APR-2000; 2000US-00560875.  
XX PD 09-AUG-2001.  
XX PP 05-FEB-2001; 2000US-00598075.  
XX PR 20-JUN-2000; 2000US-00620325.  
XX PD 09-AUG-2001.  
XX PP 05-FEB-2001; 2000US-00654936.  
XX PR 19-JUL-2000; 2000US-00663551.  
XX PD 09-AUG-2001.  
XX PP 05-FEB-2001; 2000US-00693325.  
XX PR 30-NOV-2000; 2000US-00728422.  
XX PA (HYSEQ INC.

PR 01-SEP-2000; 2000US-00654936.  
 PR 15-SEP-2000; 2000US-00664561.  
 PR 20-OCT-2000; 2000US-00693325.  
 PR 30-NOV-2000; 2000US-00728422.  
 XX  
 PA (HYSEQ INC.)  
 XX  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y; Wang ZN;  
 PI Ma QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Goodrich R;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI; 2001-476283/51.  
 DR N-PSDB; AAK52784.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 PT in diagnosis and gene therapy.  
 XX  
 PS Claim 20; Page 303; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51455-AAK53435), and the  
 CC encoded polypeptides (AAM8323-AAM8032) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK5282) and 3666 (AAM80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication.  
 XX  
 SQ Sequence 190 AA;

Query Match 100.0%; Score 754; DB 4; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-9;  
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAMVSAMSWLYLWISACAMLLCHGSLOHTFOQHHLRPEGGTCEVIAHRCNCNRRIEE 60  
 Db 58 MAMVSAMSWLYLWISACAMLLCHGSLOHTFOQHHLRPEGGTCEVIAHRCNCNRRIEE 117  
 Qy 61 RSQTVKCSCLPGKVAGTRNRPSCVDASIVGKWWCMEPCLEGECKTLPDNSGMCAT 120  
 Db 118 RSQTVKCSCLPGKVAGTRNRPSCVDASIVGKWWCMEPCLEGECKTLPDNSGMCAT 177  
 Qy 121 GNKIKITTRIHPRT 133  
 Db 178 GNKIKITTRIHPRT 190

RESULT 4  
 ABU5238  
 ID ABU5238 standard; protein; 133 AA.  
 XX  
 AC ABU5238;  
 DT 14-APR-2003 (first entry)  
 XX  
 DE Human kidney derived protein from DKFZphfkd2\_4bb6.  
 XX  
 KW Human; gene therapy; vaccine; disease treatment; detection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200112659-A2.  
 PD 22-FEB-2001.  
 XX  
 PF 18-AUG-2000; 2000WO-1B001496.  
 XX  
 PR 18-AUG-1999; 99US-0149499P.

PR 28-SEP-1999; 99US-0156503P.  
 XX  
 PA (GERHT-) GERMAN HUMAN GENOME PROJECT.  
 XX  
 PI Wiemann S;  
 XX  
 DR WPI; 2001-327840/34.  
 DR N-PSDB; ABX71301.  
 XX  
 PT Nucleic acids having the sequences of clones isolated from libraries of  
 PT different human tissues, useful in recombinant DNA methodologies.  
 XX  
 PS Claim 21; Page 426; 1095pp; English.  
 XX  
 CC This invention describes novel polynucleotides and polypeptides isolated  
 CC from human cDNA libraries which can be used for gene therapy or in  
 CC vaccines. The polynucleotides of the invention and antibodies encoded by  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate polypeptide expression. The products of the  
 CC invention may also be used to identify modulators of expression and  
 CC activity and to down regulate expression and activity. The antibodies of  
 CC the invention may also be used as diagnostic agents for detecting the  
 CC presence of polypeptides in samples. This sequence represents a  
 CC polypeptide described in the disclosure of the invention  
 XX  
 SQ Sequence 133 AA;

Query Match 98.9%; Score 746; DB 4; Length 133;  
 Best Local Similarity 99.2%; Pred. No. 1.3e-68;  
 Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAMVSAMSWLYLWISACAMLLCHGSLOHTFOQHHLRPEGGTCEVIAHRCNCNRRIEE 60  
 Db 1 MAMVSAMSWLYLWISACAMLLCHGSLOHTFOQHHLRPEGGTCEVIAHRCNCNRRIEE 60

Qy 61 RSQTVKCSCLPGKVAGTRNRPSCVDASIVGKWWCMEPCLEGECKTLPDNSGMCAT 120  
 Db 61 RSQTVKCSCLPGKVAGTRNRPSCVDASIVGKWWCMEPCLEGECKTLPDNSGMCAT 120

Qy 121 GNKIKITTRIHPRT 133  
 Db 121 GNKIKITTRIHPRT 133

RESULT 5  
 ARM3911 standard; protein; 131 AA.  
 XX  
 AC AAM39191;  
 XX  
 DE 22-OCT-2001 (First entry)  
 XX  
 DE Human polypeptide SEQ ID NO 2336.  
 XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; Shy-Drager syndrome; chemotactic;  
 KW thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 PD 26-JUL-2001.  
 XX  
 PP 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-0047175.  
 PR 21-JAN-2000; 2000US-0048825.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US0-00620312.  
 PR 03-AUG-2000; 2000US0-0053450.  
 PR 14-SEP-2000; 2000US0-0062191.  
 PR 19-OCT-2000; 2000US0-00693036.  
 XX 29-NOV-2000; 2000US0-00727344.  
 PA (HYSE-) HYSEQ INC.  
 XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; Zhou P, Goodrich R, Drmanac RT;  
 XX DR WPI; 2001-442253/47.  
 XX PT Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.  
 XX PS Example 4; SEQ ID NO 2336; 1070pp; English.  
 XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AM3842-AM4223) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemoattractant/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and CNS disorders. Note: The sequence data for this patent did not form part of the printed specification  
 XX SQ Sequence 131 AA:  
 Query Match 57.3%; Score 432; DB 4; Length 131;  
 Best Local Similarity 61.2%; Pred. No. 2e-36;  
 Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;  
 Qy 9 WVLYLWISACAMLICLHGSLQHTFOQHHLHRPESGTCBVIAAHRCNCNKRIBERSQTVKCS 68  
 Db 19 FIVTLW-----GKVSSANHHKAHHVKTGTCTEVVHLRCCNNKLEERSQTVKCS 68  
 Qy 69 CLPGKVAGTTRNPSCVDASIVIGKWWCEMPELEGECVKLDPNSGMCATGNKIKTTR 128  
 Db 69 CFPGQVAGTTRAPSCVDASIVEQKWNCHMQPLEGECKVLPDRKGWSSSGNKVKTTR 128  
 Qy 129 I 129  
 Db 129 V 129

RESULT 6  
 ID AM40977 standard; protein; 242 AA.  
 XX AC AM40977;  
 XX DT 22-OCT-2001 (first entry)  
 XX DE Human polypeptide SEQ ID NO 5908.  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW myotrophic lateral sclerosis; Shy-Drager Syndrome; chemoattractant;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX OS Homo sapiens.

XX PN WO200153312-A1.  
 XX PD 26-JUL-2001.  
 XX PF 26-DEC-2000; 2000WO-US034263.  
 XX PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-0059842.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-0066291.  
 PR 19-OCT-2000; 2000US-0069336.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; Zhou P, Goodrich R, Drmanac RT;  
 XX DR WPI; 2001-442253/47.  
 XX PT Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.  
 XX PS Example 2; SEQ ID NO 5908; 1070pp; English.  
 XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AM3842-AM4223) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemoattractant/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and CNS disorders. Note: The sequence data for this patent did not form part of the printed specification  
 XX SQ Sequence 242 AA;  
 Query Match 57.3%; Score 432; DB 4; Length 242;  
 Best Local Similarity 61.2%; Pred. No. 3.9e-36;  
 Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;  
 Qy 9 WVLYLWISACAMLICLHGSLQHTFOQHHLHRPESGTCBVIAAHRCNCNKRIBERSQTVKCS 68  
 Db 130 FIVTLW-----GKVSSANHHKAHHVKTGTCTEVVHLRCCNNKLEERSQTVKCS 179  
 Qy 69 CLPGKVAGTTRNPSCVDASIVIGKWWCEMPELEGECVKLDPNSGMCATGNKIKTTR 128  
 Db 180 CFPGQVAGTTRAPSCVDASIVEQKWNCHMQPLEGECKVLPDRKGWSSSGNKVKTTR 239  
 Qy 129 I 129  
 Db 240 V 240

RESULT 7  
 ID AAB44958 standard; protein; 121 AA.  
 XX AC AAB44958;  
 XX AC AAB44958;  
 DT 12-FEB-2001 (first entry)

XX Human secreted protein encoded by gene 42.

DE 6 AMSWVLYLWISACAMLLCHGSLOQHTFQQHHLHRPEGGTCEVIAHRCCKNRIEERSQTV 65

XX Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;

KW anti-proliferative; cytostatic; cardiotropic; cerebroprotective;

KW nootropics; neuroprotective; antibiotic; fungicide; infection;

KW ophthalmological; gene therapy; treatment; autoimmune disease;

KW cardiovascular disorder; nervous system disorder; ocular disorder;

KW wound healing; epithelial cell proliferation; skin aging; mental state;

KW transplantation; metabolism modulation.

XX Homo sapiens.

XX WO20055200-A1.

XX PD 21-SEP-2000.

XX PP 09-MAR-2000; 2000WO-US006042.

XX PR 12-MAR-1999; 99US-0124143P.

XX PR 03-DEC-1999; 99US-0168663P.

PA (HUMA ) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI; 2000-6556008/63.

DR N-PSDB; AAC79939.

XX PT Isolated human secretory proteins, nucleic acids encoding them and

PT antibodies directed against them, useful for diagnosing and treating

PT disorders related to the proteins such as cancer, Alzheimer's disease and

PT Parkinsons.

XX PS Claim 11; Page 396-397; 453pp; English.

XX CC This invention describes a novel isolated polypeptide (I) and its

CC encoding nucleic acid molecule (II) which have immunosuppressive,

CC antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiotropic,

CC vasotropic, cerebroprotective, nootropic, neuroprotective, antibacterial,

CC virucide, fungicide and ophthalmological activity and which can be used

CC for gene therapy. (I) and (II) are used to prevent, treat or ameliorate a

CC medical condition in e.g. humans, mice, rabbits, dogs, chickens, cats,

CC dogs, chickens or sheep. (I) and (II) are also used in diagnosing a

CC pathological condition or susceptibility to a pathological condition. The

CC antibodies to (II) can also be used in alleviating symptoms associated

CC with the disorders and in diagnostic immunoassays e.g. radioimmunoassays

CC or enzyme linked immunosorbent assays (ELISA). Disorders which are

CC diagnosed or treated include autoimmune diseases e.g. rheumatoid

CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or

CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular

CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders

CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and

CC fungi and ocular disorders e.g. corneal infection. The polypeptides can

CC also be used to aid wound healing and epithelial cell proliferation, to

CC prevent skin aging due to sunburn, to maintain organs before

CC transplantation, for supporting cell culture of primary tissues, to

CC regenerate tissues and in chemotaxis. The polypeptides are used to

CC modulate mammalian metabolism, to change mammal's mental state or

CC physical state by influencing biorhythms circadian rhythms depression

CC tendency for violence tolerance for pain, reproductive capabilities,

CC hormonal or endocrine levels, appetite, libido, memory, stress or other

CC cognitive qualities, as a food additive or preservative, such as to

CC increase or decrease storage capabilities, fat content, lipid, protein,

CC carbohydrates, vitamins, minerals, cofactors or other nutritional

CC components

XX SQ Sequence 121 AA;

Query Match 32.0%; Score 241; DB 3; Length 121;

Best Local Similarity 38.4%; Pred. No. 7.2e-17;

Matches 48; Conservative 23; Mismatches 42; Indels 12; Gaps 2;

XX DE 6 SMMSTFWAFMILASLLAYCS-----QLAAGTCBIVTLDRDSSQPRRTTARTA 54

XX KCSCSLPGKVAGCTTRNRPSCYDASTVIGKWCMEPCLEGBCKLDPNSGMCA-TGKNI 124

XX 55 RCACRKQIACTTRAPACDARIKTKWCDMLPCLEGBGCDLLINRSQWTCTQPGGR 114

Db Qy 125 KITTRI 129

Db Qy 115 KTTTV 119

Db

RESULT 8

AAY53016

ID AAY53016 standard; protein; 132 AA.

XX

XX Human secreted protein clone ml243\_1 protein sequence SEQ ID NO:38.

XX DE 29-FEB-2000 (First entry)

XX DT

XX Human secreted protein clone ml243\_1 protein sequence SEQ ID NO:38.

XX AC AAY53016;

XX XX

XX Human; secreted protein; nutritional; cytokine; cell proliferation;

KW differentiation; immune stimulating; growth; activin; inhibin; chemotactic;

KW haematopoiesis; regulation; tissue growth; activin; inhibin; chemotactic;

KW chemokinetic; haemostatic; thrombolytic; receptor; ligand;

KW anti-inflammatory; cadherin; tumour invasion suppressor;

KW tumour inhibition; gene therapy.

XX

XX Homo sapiens.

OS

XX WO957132-A1.

XX

XX PD 11-NOV-1999.

XX PF 07-MAY-1999; 99WO-US009970.

XX PF 07-MAY-1999; 99US-0084564P.

XX PR 02-JUN-1998; 98US-0087645P.

XX PR 22-JUL-1998; 98US-0093712P.

XX PR 31-JUL-1998; 98US-0094935P.

XX PR 10-AUG-1998; 98US-00958B0P.

XX PR 11-AUG-1998; 98US-0096689P.

XX PR 06-MAY-1999; 99US-0030611I.

XX PA (GEMY ) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Menberg D, Treacy M, Agostino MJ, Steininger RJ, Bowman MR;

XX Dibiasio-Smith E, Widom A;

XX DR WPI; 2000-052937/04.

XX DR N-PSDB; AAZ3334.

XX PT New Polymucleotides encoding secreted human proteins, derived from adult

XX placenta, adult retina, fetal brain, fetal.

XX PS Claim 47; Page 397; 492pp; English.

XX The present invention describes new human secreted proteins which were

CC isolated from adult placenta, adult retina, foetal brain, foetal kidney,

CC adult blood, adult brain, adult thyroid, adult bladder, adult neural

CC tissue, adult testes, and adult lymph node cDNA libraries. The human

CC secreted proteins, and the polynucleotides encoding them, are predicted

CC to have biological activities which would make them suitable for

CC treating, preventing or ameliorating medical conditions in humans and

CC animals. Suggested activities include nutritional activity, cytokine and

CC cell proliferation, differentiation activity, immune stimulating (e.g. as

CC vaccines) or suppressing activity, haemopoiesis regulating activity,

CC tissue growth activity, activin/inhibin activity,

CC chemokinetic/chemokine activity, haemostatic and thrombolytic activity,

CC receptor/ligand activity, adhesion/tumour

CC invasion suppressor activity, and tumour inhibition activity. The  
 CC polynucleotides are also stated to be useful for gene therapy. AA233316  
 CC to AA233373 encode human secreted proteins, and AY53060 to AY53060  
 CC represent human secreted proteins, given in the present invention  
 XX

Sequence 132 AA;

Query Match 32.0%; Score 241; DB 3; Length 132;  
 Best Local Similarity 38.4%; Pred. No. 7.9e-17;  
 Matches 48; Conservative 23; Mismatches 42; Indels 12; Gaps 2;  
 Qy 6 AMSWVLYLMSACAMLLCHGSLQHTFQOHHHLHREGGTCEVIAHRCNCNKRLEERQTV 65  
 Db 18 SMSSTFWAFMILASILAYCS-----QLAAGTCEIVLDRDSSQERTIAQTA 66  
 ID AAM24360 standard; protein; 115 AA.  
 AC XX  
 XX  
 DT 14-APR-2003 (first entry)  
 DE Human kidney-derived protein DKFZPhEkd2\_4b6 homologue.  
 XX  
 KW Human; gene therapy; vaccine; disease treatment; detection.  
 XX  
 OS Homo sapiens.  
 XX  
 EN W0200112659-A2.  
 XX  
 PD 22-FEB-2001.  
 XX  
 PR 18-AUG-2000; 2000WO-1B001496.  
 XX  
 PR 18-AUG-1999; 99US-0149499P.  
 XX  
 PA (GBHU-) GERMAN HUMAN GENOME PROJECT.  
 XX  
 PI Wiemann S;  
 XX  
 DR WPI; 2001-327840/34.

XX Nucleic acids having the sequences of clones isolated from libraries of  
 PT different human tissues, useful in recombinant DNA methodologies.  
 XX  
 PS Example III; Page 426; 105pp; English.  
 XX This invention describes novel polynucleotides and polypeptides isolated  
 CC from human cDNA libraries which can be used for gene therapy or in  
 CC vaccines. The polynucleotides of the invention and antibodies encoded by  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate polypeptide expression. The products of the  
 CC invention may also be used to identify modulators of expression and  
 CC activity and to down regulate expression and activity. The antibodies of  
 CC the invention may also be used as diagnostic agents for detecting the  
 CC presence of polypeptides in samples. This sequence represents a homologue  
 XX of a polypeptide described in the disclosure of the invention  
 Sequence 89 AA;

Query Match 31.9%; Score 240.5; DB 4; Length 89;

Best Local Similarity 49.4%; Pred. No. 5.7e-17;

Matches 44; Conservative 14; Mismatches 30; Indels 1; Gaps 1;  
 Qy 42 GTCBEVIAHRCNCNKRLEERSQVCKCCLPGKVAGTRNRSPCVDASTIVIGKRWCEWNEPC 101  
 Db 1 GTCBEVIAHRCNCNKRLEERSQVCKCCLPGKVAGTRNRSPCVDASTIVIGKRWCEWNEPC 101  
 ID AAM24360 standard; protein; 115 AA.  
 AC XX  
 XX  
 DT 12-OCT-2001 (first entry)  
 DE Human EST encoded protein SEQ ID NO: 1885.  
 XX  
 KW Human; sheep; pig; cow; fruit fly; yeast; macaque; horse;  
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
 KW diagnostics; forensic test; Gene mapping; genetic disorder; biodiversity;  
 KW gene therapy; nutrition.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200154477-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 DE 25-JAN-2001; 2001WO-US002687.  
 XX  
 PR 25-JAN-2000; 2000US-00491404.  
 PR 17-JUL-2000; 2000US-00631746.  
 PR 03-AUG-2000; 2000US-00631451.  
 PR 15-SEP-2000; 2000US-00663870.  
 XX  
 PA (HYSE-) HYSE INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
 PI Cao Y, Dumanac RA, Zhang J, Wehrman T;  
 XX  
 PR Isolated polypeptide for treatment of diseases, diagnostics, raising  
 PR antibodies and research use.  
 XX  
 PS Claim 20; Page 1217-1218; 1275pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of novel  
 CC proteins from a variety of organisms, including human, dog, cat, horse,  
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
 CC for the organism of interest. They can be used in diagnostics,  
 CC forensics, gene mapping, identification of mutations, to assess  
 CC biodiversity and for nutritional purposes. The present sequence is a  
 XX protein of the invention  
 Sequence 115 AA;  
 XX  
 PR Query Match 29.4%; Score 221.5; DB 4; Length 115;  
 CC Best Local Similarity 54.0%; Pred. No. 6.8e-15;  
 CC Matches 47; Conservative 6; Mismatches 21; Indels 13; Gaps 2;  
 CC  
 Qy 5 SAMSWVL----YLWISACAMLLCHGSLQHTFQOHHLRPEGGTCAVIAHRCNCNKRLE 59  
 Db 10 STGGWLLALCLAWLWTHLTLAALQPTATLVQQ----GTCEVIAHRCNCNKRLE 61  
 Qy 60 ERSQTYKCSCLPGKVAGTTNRPSCVD 86



Db 419 CHGQCOH-----GGTCKDLVNGYQCVCPRGFGRRHELEYKCASSPCRRGGIC 467  
 Qy 59 EERSQTIVKCSCLPGRVAGTRNPRSC-VDASIVIGRWCENEPCLEBECBCTLPDNGWM 117  
 :  
 Db 468 EDLVDFGRCHC-PRGLSG----PLCEVDVDL---WCEPNPCLNGARCYNLEDD--YY 514  
 Qy 118 QA 119  
 Db 515 CA 516

RESULT 13  
 ID AAW95557 standard; protein; 1964 AA.  
 XX AC AAW95557;  
 XX DT 16-JUN-1999 (first entry)  
 XX DE Mus musculus notch4 protein.  
 XX KW notch4; int-3; antisense; angiogenesis; modulation; inhibition;  
 KW promotion; solid tumour; haemangioma; haemangiosarcoma; Kaposi's sarcoma;  
 KW ischaemia; gangrene; treatment; hemangioma; diabetic ulcers;  
 KW chronic ulceration; hemangiosarcoma;  
 KW cerebral autosomal dominant arteriopathy; subcortical infarction;  
 KW leucoencephalopathy; vascular dementia; wound healing.

OS Mus musculus.  
 XX Key Location/Qualifiers  
 FH Region 804 .806 /note= "putative glycosylation site"  
 FT Region 1139 .1141 /note= "putative glycosylation site"  
 FT Region 1411 /note= "int-3 oncoprotein initiating Methionine"  
 FT Domain 1441 .1164 /note= "transmembrane domain"  
 FT Domain 1925 .1941 /note= "putative PEST domain"  
 FT .  
 PN WO9857621-A1.  
 XX PD 23-DEC-1998.  
 XX PF 18-JUN-1998; 98WO-US013050.  
 XX PR 18-JUN-1997; 97US-00879351.  
 XX PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX PI Kitajewski, J; Uttenraad, H;  
 XX DR WPI; 1999-095286/08.  
 XX DR N-PSDB; AXX00989.

PT New nucleic acid encoding Notch4 protein - used to modulate angiogenesis,  
 e.g. for treatment of tumours, ischaemia, ulcers and wounds

CC The sequence is that of the notch4 protein. Antibodies (Ab) raised  
 CC against it are used to block binding of notch4 protein to its ligand.  
 CC Agonists and antagonists are used to modulate (promote or inhibit,  
 CC respectively) angiogenesis; this is also promoted by transducing cells to  
 CC express notch4, and optionally its ligand, or inhibited by treatment with  
 CC Ab, or their fragments. Specifically inhibition of angiogenesis is used  
 CC to treat solid tumours, haemangiomas, haemangiosarcoma and Kaposi's  
 CC sarcoma, while its promotion is used to treat ischaemia, gangrene,  
 CC diabetic ulcers, chronic ulceration, cerebral, autosomal dominant  
 CC arteriopathy with subcortical infarcts and leucoencephalopathy, vascular  
 CC dementia and wounds. Ab are also used to detect notch4 protein on the

CC surface of cells. Signalling through notch4: (a) is distinct from that  
 CC through fibroblast growth factor (FGF), so may effect angiogenesis not  
 CC responsive to FGF modulation; (b) determines the fate of a cell, so is  
 CC effective in presence of other angiogenic factors, and (c) may show a  
 CC synergistic effect with FGF signalling

XX Sequence 1964 AA;

Query Match Score 96; DB 2; Length 1964;  
 Best Local Similarity 26.8%; Pred. No. 1.2;  
 Matches 33; Conservative 8; Mismatches 48; Indels 34; Gaps 6;

Qy 15 ISACAMLLC-HGSLOHTFQOHHLHRPEGGTC-----EVIAAHR--CCNKRI 58  
 DB 689 LGGCSTPCAGGTCP-----QFSGYNTCPAGMGTCSSEVTACHSGPCLNGGSC 741  
 Qy 59 EERSQTIVKCSCLPGRVAGTRNRPSCDASIVIGKWWCMEPCLGECECKTLPDNSGMC 118  
 DB 742 SIRPEGYSCTRLP---SHTGRHCQTAVD-----HCVSASCLNGGTCVNKPGTFFCLC 790  
 Qy 119 ATG 121  
 DB 791 ATG 793

RESULT 14  
 ID ADE63713 standard; protein; 2531 AA.  
 XX AC ADE63713;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Rat Protein CRA40667, SEQ ID NO 9657.  
 XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; cci; spared nerve injury; SNL; Chung.  
 XX OS Rattus norvegicus.  
 XX PN WO2003016475-A2.  
 XX PD 27-FEB-2003.  
 XX PF 14-AUG-2002; 2002NO-US025765.  
 XX PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 PA (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX Woolf, C; D'urso, D; Befort, K; Costigan, M;  
 XX DR WPI; 2003-268312/26.  
 XX GENBANK; CAA40667.  
 PT New composition comprising two or more isolated polypeptides, useful for  
 preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 claimed are a vector comprising the novel polynucleotide, a host cell  
 comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially expressed in an animal subjected to pain and a  
 kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates

XX

CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC the sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic form directly from WIPO at  
 CC [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences).

XX Sequence 2531 AA;

Query Match Score 95.5; DB 7; Length 2531;  
 Best Local Similarity 12.7%; Pred. No. 1.7;  
 Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;  
 Qy 15 ISACAM1ICHGSLOQHTFQQHHLHRPEGCTCE-VIAAHC-----ONKR 56  
 Db 679 IDECAGSPH-----NGTCTEDJAGFTCRCPGYHDITCLSSVNECNSN 723  
 Qy 57 ---RIBERSQTVKCSCLPGLKYGATGTPRNPSCVDASTIVGKWCMEPCLSGECKTLP 111  
 Db 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVNCGTCKDM- 771  
 Qy 112 DNGWMC 118  
 Db 772 -TSGYVC 777

**RESULT 15**  
 ID ADB63705 standard; protein; 2531 AA.  
 XX ADB63705;  
 AC :  
 DT 29-JAN-2004 (first entry)  
 XX DE Rat Protein CAA40667, SEQ ID NO 9649.  
 XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
 XX OS Rattus norvegicus.  
 XX PN WO2003016475-A2.  
 XX PD 27-FEB-2003.  
 XX PF 14-AUG-2002; 2003WO-US025755.  
 XX PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX PA (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.

XX WPI: 2003-268312/26.  
 DR GENBANK; CAR0667.  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX

CC Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC the sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC [ftp://wipo.int/pub/published\\_pct\\_sequences](ftp://wipo.int/pub/published_pct_sequences).

CC SQ Sequence 2531 AA;  
 CC Query Match Score 95.5; DB 7; Length 2531;  
 CC Best Local Similarity 12.7%; Pred. No. 1.7;  
 CC Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;  
 CC Qy 15 ISACAM1ICHGSLOQHTFQQHHLHRPEGCTCE-VIAAHC-----ONKR 56  
 CC Db 679 IDECAGSPH-----NGTCTEDJAGFTCRCPGYHDITCLSSVNECNSN 723  
 CC Qy 57 ---RIBERSQTVKCSCLPGLKYGATGTPRNPSCVDASTIVGKWCMEPCLSGECKTLP 111  
 CC Db 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVNCGTCKDM- 771  
 CC SQ Sequence 2531 AA;  
 CC Query Match Score 95.5; DB 7; Length 2531;  
 CC Best Local Similarity 12.7%; Pred. No. 1.7;  
 CC Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;  
 CC Qy 15 ISACAM1ICHGSLOQHTFQQHHLHRPEGCTCE-VIAAHC-----ONKR 56  
 CC Db 679 IDECAGSPH-----NGTCTEDJAGFTCRCPGYHDITCLSSVNECNSN 723  
 CC Qy 57 ---RIBERSQTVKCSCLPGLKYGATGTPRNPSCVDASTIVGKWCMEPCLSGECKTLP 111  
 CC Db 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVNCGTCKDM- 771  
 CC SQ Sequence 2531 AA;

Search completed: April 8, 2004, 13:16:11  
 Job time : 58 secs

Result No.	Score	Query Match Length	DB ID	Description	
1	11.0	14.7	JAG2_RAT	P97607 rattus norvegicus protein; JAG2; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
2	9.6	12.7	NTC4_MOUSE	P31695 mus musculus protein; NTC4; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
3	95.5	12.7	NTCL_RAT	Q01705 mus musculus protein; NTCL; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
4	9.5	12.7	NTCL_RAT	Q01708 rattus norvegicus protein; NTCL; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
5	92.5	12.3	CRR_DROME	P10040 drosophila melanogaster protein; CRR; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
6	91.5	12.1	JAG3_BRARE	Q9Y54 brachydanio rerio protein; JAG3; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
7	12.0	5.5	DICDI	P04503 dictyostelia intestinalis protein; DICDI; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
8	86.5	11.5	NOTC_DROME	P07207 drosophila melanogaster protein; NOTC; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
9	86	11.4	685	JAG1_HUMAN	Q9n61 homo sapiens protein; JAG1; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
10	86	11.4	1218	JAG1_MOUSE	Q9xxk0 mus musculus protein; JAG1; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
11	86	11.4	1218	JAG1_HUMAN	Q63722 rattus norvegicus protein; JAG1; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
12	85.5	11.3	1064	FBP1_STRONGYL	P10079 strongyloma vulgaris protein; FBP1; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
13	84.5	11.2	2003	NTC4_HUMAN	Q99466 homo sapiens protein; NTC4; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
14	84	11.1	2247	JAG2_MOUSE	Q9yqes mus musculus protein; JAG2; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
15	84	11.1	2524	NOTC_XENLA	P21783 xenopus laevis protein; NOTC; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
16	83.5	11.1	1170	TSP2_BOVIN	Q95116 bos taurus protein; TSP2; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
17	82.5	10.9	NTC2_MOUSE	Q35516 mus musculus protein; NTC2; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.	
18	82.5	10.9	2471	NTC2_HUMAN	Q04721 homo sapiens protein; NTC2; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
19	82.5	10.9	2471	NTC2_RAT	Q9Tw30 rattus norvegicus protein; NTC2; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
20	80.5	10.9	2471	NTC3_MOUSE	Q61982 mus musculus protein; NTC3; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
21	82	10.9	2318	NTC3_HUMAN	Q07954 homo sapiens protein; NTC3; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
22	82	10.9	4554	ZAN_MOUSE	Q88799 mus musculus protein; ZAN; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
23	82	10.9	5376	JAG1_HUMAN	Q04018 saccharomyces cerevisiae protein; JAG1; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
24	81.5	10.8	355	YMT5_YEAST	Q94018 saccharomyces cerevisiae protein; YMT5; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
25	81.5	10.8	2556	NTC3_HUMAN	Q9um47 homo sapiens protein; NTC3; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
26	80.5	10.7	2321	NTC3_HUMAN	Q9um47 homo sapiens protein; NTC3; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
27	80	10.6	937	VME_BOVIN	P80012 bos taurus protein; VME; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
28	79.5	10.5	2319	NTC3_RAT	Q9r172 rattus norvegicus protein; NTC3; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
29	79	10.5	819	AD09_HUMAN	Q13443 homo sapiens protein; AD09; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
30	78.5	10.4	1178	TSP2_CHICK	P35440 gallus gallus protein; TSP2; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
31	78	10.4	3312	CLR3_HUMAN	Q9nyq7 homo sapiens protein; CLR3; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
32	78	10.3	1191	LRG2_MOUSE	Q61092 mus musculus protein; LRG2; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.
33	78	10.3	1376	CRBB_HUMAN	P82279 homo sapiens protein; CRBB; LOCATED IN BRAIN; SUBCELLULAR LOCATION: Type I membrane protein.

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 141681 seqs, 52070155 residues  
Total number of hits satisfying chosen parameters: 141681  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42::\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

%

RESULT 1

JAG2\_RAT

ID: P97607;  
AC: P97607;  
DT: 28-FEB-2003 (Rel. 41, Created)  
DT: 28-FEB-2003 (Rel. 41, Last sequence update)  
DT: 15-MAR-2004 (Rel. 41, Last annotation update)

Jagged2 (Jagged2) (Fragment).

CN: JAG2.

OS: Rattus norvegicus (Rat).

RA: Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

[1]

CC: FUNCTION: Putative Notch ligand involved in the mediation of Notch signaling. May have a role in neurogenesis in the peripheral nervous system, limb development and in the adult brain.

CC: -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC: -!- DEVELOPMENTAL STAGE: At stage E12.5 it is detected in dorsal root ganglia, AER, and surface ectoderm. At E14.5, found as well in cranial ganglia, thymus and olfactory epithelia. At E16.5, found as well in salivary gland, tooth buds and hair follicles.

CC: -!- SIMILARITY: Contains 16 EGF-like domains.

CC: -!- SIMILARITY: Contains 1 DSL domain.

CC: -!- SIMILARITY: Contains 1 DSL domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC: EMBL: U70050; AAC52946.1; -

DR: HSSP; P00743; 1CCF.

DR: GO; GO:000912; P:cell fate determination; NAS.

DR: GO; GO:0008083; P:growth factor activity; ISS.

DR: GO; GO:000512; P:Notch binding; NAS.

DR: GO; GO:0007605; P:cell cycle; ISS.

DR: GO; GO:0007049; P:cell differentiation; NAS.

DR: GO; GO:00020154; P:cell fate commitment; ISS.

DR: GO; GO:0001709; P:cell fate determination; NAS.

DR: GO; GO:000912; P:hair cell fate commitment; ISS.

DR: GO; GO:0007605; P:hair cell fate commitment; ISS.

DR: GO; GO:0003036; P:limb morphogenesis; NAS.

DR: GO; GO:0003219; P:N signaling pathway; NAS.

DR: GO; GO:0003034; P:regulation of cell migration; ISS.

DR: GO; GO:0042127; P:regulation of cell proliferation; ISS.

DR: GO; GO:0010283; P:spemato genesis; IEP.

DR: GO; GO:0030217; P:T-cell differentiation; ISS.

DR: GO; GO:0045061; P:thymic T-cell selection; ISS.



RA Rowen L., Mahairas G., Qin S., Ahearn M.E., Dankers C., Lasky S.,  
 RA Lorez C., Schmidt S., Tipton S., Traicoff R., Zackrone K., Hood L.,  
 RT "Sequence of the mouse major histocompatibility locus class III  
 region.";  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.  
 [5]

RN SEQUENCE OF 1436-1600 FROM N A.  
 RP MEDLINE=9952211; PubMed=1023982;  
 RA Lee J.-S., Haruna T., Ishimoto A., Honjo T., Yanagawa S.-I., Notch4/int3  
 RT "Intracisternal type A particle-mediated activation of truncated Notch4/int3  
 gene in a mouse mammary tumor: generation of retroviral splicing events.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:5643-5648 (2001).  
 J. Virol. 73:5166-5171(1999).

[6]

RP SEQUENCE OF 1463-1964, POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS  
 OF VAL-1453.  
 RP MEDLINE=21244657; PubMed=11344305;  
 RX UTtendele H., Ho J., Rossant J., Kitajewski J.;  
 RA Saxena M.T., Schroeter E.H., Mumt J.S., Kopan R.;  
 RT "Murine notch homologs (N1-4) undergo presenilin-dependent  
 proteolysis.";  
 RT Notch4 in embryonic endothelium.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:5643-5648 (2001).  
 J. Biol. Chem. 276:40268-40273 (2001).  
 RP POST-TRANSLATIONAL PROCESSING.  
 RX MEDLINE=21374376; PubMed=11459941;

RA Mizurani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;  
 RT "Conservation of the biochemical mechanisms of signal transduction  
 among mammalian Notch family members.";  
 Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031 (2001).  
 RL

[8]

CC JAGged2, Jagged2 and Deltal to regulate cell-fate determination.  
 upon ligand activation through the released notch intracellular  
 domain (NICD), it forms a transcriptional activator complex with  
 RBP-J kappa and activates genes of the enhancer of split locus.  
 Affects the implementation of differentiation, proliferation and  
 apoptosis (By similarity). May regulate branching  
 morphogenesis in the developing vascular system.

CC SUBUNIT: Heterodimer of a C-terminal fragment N (TM) and a N-  
 terminal fragment N (EC) which are probably linked by disulfide  
 bonds.

CC SUBCELLULAR LOCATION: Type I membrane protein. Following  
 proteolytic processing NICD is translocated to the nucleus.

CC TISSUE SPECIFICITY: Highly expressed in lung, moderately in heart,  
 kidney, and at lower levels in the ovary and skeletal muscle. A  
 very low expression is seen in the brain, intestine, liver and  
 testis.

CC DEVELOPMENTAL STAGE: Highly expressed in endothelial cells during  
 embryonic development from 9.0 dpc.

CC PTM: Synthesized in the endoplasmic reticulum as an inactive form  
 which is proteolytically cleaved by a furin-like convertase in the  
 trans-Golgi network before it reaches the plasma membrane to yield  
 an active, ligand-accessible form. Cleavage results in a C-  
 terminal fragment N (TM) and a N-terminal fragment N (EC). Following  
 ligand binding, it is cleaved by TNF-alpha converting enzyme  
 (TACE) to yield a membrane-associated intermediate fragment called  
 notch extracellular truncation (NEXT). This fragment is then  
 cleaved by presenilin dependent gamma-secretase to release a  
 notch-derived peptide containing the intracellular domain (NICD)  
 from the membrane.

CC PTM: Phosphorylated.

CC DISEASE: Loss of the extracellular domain causes constitutive  
 activation of the Notch protein, which leads to hyperproliferation  
 of glandular epithelial tissues and development of mammary  
 carcinomas.

CC SIMILARITY: Belongs to the NOTCH family.

CC SIMILARITY: Contains 29 EGF-like domains.

CC SIMILARITY: Contains 3 Lin/Notch repeats.

-!- SIMILARITY: Contains 5 ANK repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).  
 CC DR EMBL; M80456; AAB8377..1;  
 CC DR EMBL; U43631; AAC52630..1;  
 CC DR EMBL; U43691; AAC52631..1;  
 CC DR EMBL; AF030001; AAB82004..1;  
 CC DR EMBL; AB016771; BAA32281..1; ALT SEQ.  
 CC DR EMBL; AB016772; BAA32283..1; ALT INIT.  
 CC DR EMBL; AB016773; BAA32284..1; ALT INIT.  
 CC DR PIR; A38072; TMVNT3.  
 CC DR PIR; T09059; T09059.  
 CC DR HSSP; P08709; 1BF9.  
 CC DR MGD; MGI:1.07473; Notch4.  
 CC DR InterPro; IPR002110; ANK.  
 CC DR InterPro; IPR000152; ASX\_hydroxyl\_S.  
 CC DR InterPro; IPR000742; EGF\_2.  
 CC DR InterPro; IPR001881; EGF\_Ca.  
 CC DR InterPro; IPR001438; EGF\_I.  
 CC DR InterPro; IPR006209; EGF\_Like.  
 CC DR InterPro; IPR002049; Lammin\_EGF.  
 CC DR InterPro; IPR008209; Notch.  
 CC DR InterPro; IPR008080; Notch\_dom.  
 CC PFam; PF00023; ank; 6.  
 CC DR SMART; SM00179; EGF; 27.  
 CC PFam; PF00066; notch; 2.  
 CC PIRSF; PIRSF002279; Notch; 1.  
 CC PRINTS; PR00010; EGF\_BLOOD.  
 CC PRINTS; PR00011; EGF\_LAMININ.  
 CC PRINTS; PR01452; NOTCH.  
 CC DR SMART; SM00248; ANK; 6.  
 CC DR PROSITE; PS00022; EGF\_1;  
 CC DR PROSITE; PS00022; EGF\_2;  
 CC DR PROSITE; PS00025; EGF\_3;  
 CC DR PROSITE; PS01187; EGF\_CA; 9.  
 CC KW Receptor; transcription regulation; Activator; Differentiation;  
 CC KW Developmental Protein; Repeat; ANK repeat; EGF-like domain;  
 CC KW Transmembrane; Glycoprotein; Signal; Phosphorylation; Proto-oncogene.  
 CC FT SIGNAL 1 20  
 CC FT CHAIN 21 1964  
 CC FT PROSITE; PS00022; EGF\_1;  
 CC FT PROSITE; PS00022; EGF\_2;  
 CC FT PROSITE; PS00025; EGF\_3;  
 CC FT PROSITE; PS01187; EGF\_CA; 9.  
 CC FT TRANSMEM 1444 1464  
 CC FT DOMAIN 1465 1964  
 CC FT CHAIN 1411 1964  
 CC FT DOMAIN 1428 1964  
 CC FT DOMAIN 1463 1964  
 CC FT DOMAIN 1443 1964  
 CC FT SIGNAL 1 20  
 CC FT DOMAIN 21 1964  
 CC FT DOMAIN 231 271  
 CC FT DOMAIN 61 112  
 CC FT DOMAIN 115 152  
 CC FT DOMAIN 153 189  
 CC FT DOMAIN 191 229  
 CC FT DOMAIN 231 271  
 CC FT DOMAIN 273 309  
 CC FT DOMAIN 311 350  
 CC FT DOMAIN 352 388  
 CC FT DOMAIN 389 427  
 CC FT DOMAIN 429 470  
 CC FT DOMAIN 472 508  
 CC FT DOMAIN 510 546  
 CC FT DOMAIN 548 584  
 CC



-!- ALTERNATIVE PRODUCTS:  
Event-Alternative splicing; Named isoforms=2;  
Name=";  
IsoId=Q01705\_1; Sequence=Displayed;

CC Name=";  
IsoId=Q01705\_2; Sequence=VSP\_001402, VSP\_001403, VSP\_001404;  
Note="No experimental confirmation available";  
TISSUE\_SPECIFICITY: Highly expressed in the brain, lung and thymus. Expressed at lower levels in the spleen, bone-marrow, spinal cord, eyes, mammary gland, liver, intestine, skeletal muscle, kidney and heart.

CC DEVELOPMENTAL\_STAGE: First detected in the mesoderm at 7.5 dpc. By 8.5 dpc highly expressed in presomitic mesoderm, mesenchyme and endothelial cells, while much lower levels are seen in the neuroepithelium. Between 9.5-10.5 dpc expressed at high levels in the neuroepithelium. At 13.5 dpc expressed in the surface ectoderm, eye and developing whisker follicles.

CC PTM: Synthesized in the endoplasmic reticulum as an inactive form which is proteolytically cleaved by a furin-like convertase in the trans-Golgi network before it reaches the plasma membrane to yield an active, ligand-accessible form. Cleavage results in a C-terminal fragment N(TM) and a N-terminal fragment N(EC). Following ligand binding, it is cleaved by TNF alpha converting enzyme (TACE) to yield a membrane-associated intermediate fragment called notch extracellular truncation (NEXT). This fragment is then cleaved by presenilin dependent gamma-secretase to release a notch-derived peptide containing the intracellular domain (NICD) from the membrane.

CC PTM: Phosphorylated.

CC SIMILARITY: Belongs to the NOTCH family.  
CC SIMILARITY: Contains 36 EGF-like domains.  
CC SIMILARITY: Contains 3 Lin-Notch repeats.  
CC SIMILARITY: Contains 5 ANK repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC MGD; MGI:97363; Notch1.  
DR GO; GO:0005887; C:integral to plasma membrane; IC.  
DR GO; GO:000515; F:protein binding; IPI.  
DR GO; GO:0030154; P:cell differentiation; IMP.  
DR GO; GO:0007386; P:compartment specification; IMP.  
DR GO; GO:0007219; P:N signaling pathway; IC.  
DR GO; GO:045944; P:positive regulation of transcription from P...; IDA.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000152; Asx\_hydroxyly\_S.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001438; EGF\_II.  
DR InterPro; IPR006209; EGF\_11like.  
DR InterPro; IPR002110; Laminin\_EGF.  
DR InterPro; IPR008297; Notch\_Notch\_dom.  
DR Pfam; PF00023; ank; 7.  
DR Pfam; PF00008; EGF\_35.  
DR Pfam; PF00666; notch; 3.  
DR Pfam; PF002273; notch; 1.  
DR Prints; PR00010; EGFBLOOD.  
DR Prints; PR00011; EGFLAMININ.  
DR Prints; PR01452; NOTCH.

CC SMART; SM00248; ANK; 6.  
SMART; SM00119; EGF\_Ca; 24.  
DR SMART; SM00094; NL\_2.  
DR PROSITE; PS50297; ANK REP REGION; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 2.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 22.  
DR PROSITE; PS00022; EGF\_1; 34.  
DR PROSITE; PS01186; EGF\_2; 27.  
DR PROSITE; PS50026; EGF\_3; 36.  
KW Receptor; transcription regulation; Activator; Differentiation; Developmental protein; Repeat; ANK repeat; EGF-like domain; Transmembrane; Glycoprotein; Signal; Phosphorylation; Alternative splicing.  
KW POTENTIAL.  
FT SIGNAL 1 18  
FT CHAIN 19 2531  
FT CHAIN 1711 2531  
FT CHAIN 1744 2531  
FT DOMAIN 19 1725  
Query Match 12.7%; Score 95.5; DB 1; Length 2531;  
Best Local Similarity 26.8%; Pred. No. 0.2;  
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;  
Oy 15 ISACAMLCHSLQHFTQQHHLRPGGTCB-VIAHRC-----CNGN 56  
Db 679 IDECAGSPCH-----NGCTEDGAGFTCPCEGYHDPTCLSEVNENCSN 723  
Db 724 PCIRGACRDGLNGYKCDAPS-WSGTNCD-----INNNNECESNPVCYNGGTCKDM- 771  
Oy 112 DNSGMNC 118  
Db 772 -TSGYTC 777  
RESULT 4  
ID NTIC1\_RAT STANDARD;  
AC Q07056;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Neurogenic locus notch precursor 1 precursor (Notch 1).  
GN NOTCH.  
Ox Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Rattus.  
OX NCBI\_TAXID=10116;  
RN [1] SEQUENCE FROM N\_A.  
RP TISSUE-Schwann cell;  
RX MEDLINE=92111383; PubMed=1764995;  
RA Weinmaster G.; Roberts V.J.; Lemke G.;  
RT "A homolog of Drosophila Notch expressed during mammalian development";  
RL Development 113:199-205 (1991).  
RN [2] REVISIONS TO 1652-1653.  
RA Weinmaster G.;  
RN submitted (APR-1998) to the EMBL/GenBank/DDBJ databases  
[3] FUNCTION.  
RP MEDLINE=21034508; PubMed=11182080;  
RX Tanigaki K.; Nozaki F.; Takahashi J.; Tashiro K.; Kurooka H.;  
RA Honjo T.;  
RT "Notch1 and Notch3 instructively restrict bFGF-responsive multipotent neural progenitor cells to an astroglial fate";  
RL Neuron 29:45-55(2001).  
RN [4] TISSUE SPECIFICITY.  
RP MEDLINE=93202015; PubMed=1295745;



FT	DOMAIN	2497	2500	POLY-SER.	CLEAVAGE BY (FURIN-LIKE PROTEASE)	(BY
FT	SITE	1654	1655		SIMILARITY.	
FT	DISULFID	24	37		BY SIMILARITY.	
FT	DISULFID	31	46		BY SIMILARITY.	
FT	DISULFID	48	57		BY SIMILARITY.	
FT	DISULFID	63	74		BY SIMILARITY.	
FT	DISULFID	68	87		BY SIMILARITY.	
FT	DISULFID	89	98		BY SIMILARITY.	
FT	DISULFID	106	117		BY SIMILARITY.	
FT	DISULFID	111	127		BY SIMILARITY.	
FT	DISULFID	129	138		BY SIMILARITY.	
FT	DISULFID	144	155		BY SIMILARITY.	
FT	DISULFID	149	164		BY SIMILARITY.	
FT	DISULFID	166	175		BY SIMILARITY.	
FT	DISULFID	182	195		BY SIMILARITY.	
FT	DISULFID	189	204		BY SIMILARITY.	
FT	DISULFID	206	215		BY SIMILARITY.	
FT	DISULFID	222	233		BY SIMILARITY.	
FT	DISULFID	227	243		BY SIMILARITY.	
Query Match Score	95.5;	DB 1;	Length 2531;			
Best Local Similarity	26.8%;	Pred. No. 0.2;				
Matches 34;	Conservative 7;	Mismatches 35;	Indels 51;	Gaps 7;		
Qy	15	ISACAMILCHGSTQHTFQOHHLRPEGGTCE-VIAAHR-		-CNKN 56		
Db	679	1DECAGSICH-----NGGTCEDGJAGFTCRPGYHDPDTCLSVNNECNSN	723			
Qy	57	--RIBERSQTQYKCSCLPQKGKVAGTTPRNPSVCDASIVGKWCMEPCPLGECKTLP	111			
Db	724	PCIHGACRDGLNGYKCDAPG--WSGTNCD-----INNNECESNPVCNNGTCKDM-	771			
Qy	112	DNSGWMC 118				
Db	772	-TSGYVC 777				
RESULT	5					
CRB_DRONE	ID	PR00011; EGFBL000D				
STANDARD;						
AC	PR00040;					
DT	01-MAR-1989 (Rel. 10, Created)					
DT	01-MAY-1991 (Rel. 18, Last sequence update)					
DT	28-FEB-2003 (Rel. 41, Last annotation update)					
DE	Crumbs protein precursor (95F).					
GN	Drosophila melanogaster (Fruit Fly)					
OS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;					
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;					
OC	Ephydriodea; Drosophilidae; Drosophila.					
OX	NCBI_TAXID=7227;					
RN	SEQUENCE FROM N.A.					
RC	SPRAIN-Oregon-R; TISSUE=Embryo;					
RX	MEDLINE-90263104; Pubmed=244615;					
RA	Topass U., Theres C.; Knust E.;					
RT	"Crumbs encodes an EGF-like protein expressed on apical membranes of Drosophila epithelial cells and required for organization of epithelia."					
RT	Cell 61:787-799(1990).					
RN	SEQUENCE OF 1663-1955 FROM N.A.					
RC	TISSUE=Embryo;					
RX	Medline=87216537; PubMed=310986;					
RA	Knust E., Dietrich U., Topass U., Bremer K.A., Weigel D., Vaessin H., Campos-Ortega J.A.;					
RA	"EGF homologous sequences encoded in the genome of Drosophila melanogaster, and their relation to neurogenic genes";					
RT	EMBO J. 6:761-766(1987).					
RL	FUNCTION: May play a role in the development of epithelia, possibly for the establishment and/or maintenance of cell polarity. It may act as a signal.					
CC						
CC						

|- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC |- PTM: PHOSPHORYLATED IN THE CYTOPLASMIC DOMAIN (POTENTIAL).  
 CC |- !- SIMILARITY: Contains 3 laminin G-like domains.  
 CC |- !- SIMILARITY: Contains 3 laminin G-like domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions. This entry is as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC M33753; AAA2B828-1; ALT SEQ.  
 CC EMBL; X05144; CRA28793.1; -.  
 DR PIR; A35672; A35672.  
 DR PIR; B26637; B26637.  
 DR HSSP; P00740; 1EDM.  
 DR FlyBase; FBgn0000368; crb.  
 DR GO; GO\_0016324; C:apical plasma membrane; IDA.  
 DR GO; GO\_0016327; C:apical-lateral plasma membrane; IDA.  
 DR GO; GO\_0007163; P:establishment and/or maintenance of cell polarity; IMP.  
 DR GO; GO\_0016332; P:establishment and/or maintenance of polarit.  
 DR GO; GO\_0016334; P:establishment and/or maintenance of polarit.; IMP.  
 DR GO; GO\_0045194; P:photoreceptor maintenance; IMP.  
 DR GO; GO\_0042052; P:rhabdomere development; NAS.  
 DR GO; GO\_0045186; P:zonula adherens assembly; IMP.  
 DR InterPro; IPR00152; Asx hydroxyl S.  
 DR InterPro; IPR00985; Connexin\_lec\_gli.  
 DR InterPro; IPR00742; EGF\_Z.  
 DR InterPro; IPR00438; EGF\_T1.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR InterPro; IPR003049; Laminin\_EGF.  
 DR InterPro; IPR001791; Laminin\_G.  
 Pfam; PF00008; EGF; 26.  
 Pfam; PF00054; Laminin\_G; 3.  
 PRINTS; PR00010; EGFBL000D  
 PRINTS; PR00011; EGFAMININ.  
 SMART; SM00179; EGF CA; 11.  
 SMART; SM00282; LamG; 3.  
 PROSITE; PS00010; ASX\_HYDROXYL; 15.  
 PROSITE; PS00022; EGF\_F; 26.  
 PROSITE; PS50066; EGF\_3.; 27.  
 PROSITE; PS00187; EGF CA; 12.  
 PROSITE; PS00025; LAM G\_DOMAIN; 3.  
 KW Differentiation; Repeat; EGF-like domain; Transmembrane;  
 KW Glycoprotein; Signal; Phosphorylation.  
 FT SIGNAL; 1 90 CRUMBS PROTEIN.  
 FT DOMAIN 91 2139 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 2085 2084 POTENTIAL.  
 FT DOMAIN 2112 2111 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 267 2139 EGF-LIKE 1.  
 FT DOMAIN 303 303 EGF-LIKE 2.  
 FT DOMAIN 348 348 EGF-LIKE 3.  
 FT DOMAIN 388 425 EGF-LIKE 4; CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 427 463 EGF-LIKE 5.  
 FT DOMAIN 464 500 EGF-LIKE 6.  
 FT DOMAIN 501 532 EGF-LIKE 7.  
 FT DOMAIN 545 581 EGF-LIKE 8.  
 FT DOMAIN 582 611 EGF-LIKE 9.  
 FT DOMAIN 609 646 CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 648 685 EGF-LIKE 11.  
 FT DOMAIN 687 723 CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 725 761 CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 763 800 CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 802 838 EGF-LIKE 15.  
 FT DOMAIN 840 902 EGF-LIKE 16.  
 FT DOMAIN 904 940 EGF-LIKE 17; CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 942 942 EGF-LIKE 18; CALCIUM-BINDING (POTENTIAL).

980	1021	BFGF-LIKE 19.	FT	DISULFID	1233	1242	BY SIMILARITY.
DOMAIN	1023	LAMININ G-LIKE 1.	FT	DISULFID	1485	1496	BY SIMILARITY.
DOMAIN	1203	BFGF-LIKE 20.	FT	DISULFID	1490	1505	BY SIMILARITY.
DOMAIN	1207	LAMININ G-LIKE 2.	FT	DISULFID	1507	1516	BY SIMILARITY.
DOMAIN	1243	BFGF-LIKE 21.	FT	DISULFID	1763	1774	BY SIMILARITY.
DOMAIN	1250	LAMININ G-LIKE 3.	FT	DISULFID	1768	1783	BY SIMILARITY.
DOMAIN	1480	BFGF-LIKE 22.	FT	DISULFID	1785	1794	BY SIMILARITY.
DOMAIN	1481	BFGF-LIKE 23.	FT	DISULFID	1812	1821	BY SIMILARITY.
DOMAIN	1558	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1806	1821	BY SIMILARITY.
DOMAIN	1758	BFGF-LIKE 24.	FT	DISULFID	1823	1832	BY SIMILARITY.
DOMAIN	1759	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1839	1850	BY SIMILARITY.
DOMAIN	1795	BFGF-LIKE 25.	FT	DISULFID	1844	1859	BY SIMILARITY.
DOMAIN	1797	BFGF-LIKE 26.	FT	DISULFID	1861	1870	BY SIMILARITY.
DOMAIN	1833	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1878	1889	BY SIMILARITY.
DOMAIN	1835	BFGF-LIKE 27.	FT	DISULFID	1883	1903	BY SIMILARITY.
DOMAIN	1871	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1905	1914	BY SIMILARITY.
DOMAIN	1915	BFGF-LIKE 28.	FT	DISULFID	1919	1930	BY SIMILARITY.
DOMAIN	1915	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1924	1939	BY SIMILARITY.
DOMAIN	1953	BFGF-LIKE 29.	FT	DISULFID	1941	1950	BY SIMILARITY.
DOMAIN	1989	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1957	1968	BY SIMILARITY.
DOMAIN	2029	BFGF-LIKE 29.	FT	DISULFID	1977	1977	BY SIMILARITY.
DOMAIN	2070	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	1984	1984	BY SIMILARITY.
DISULFID	271	BFGF-LIKE 29.	FT	DISULFID	1985	1985	BY SIMILARITY.
DISULFID	276	BY SIMILARITY.	FT	DISULFID	1986	1986	BY SIMILARITY.
DISULFID	293	BY SIMILARITY.	FT	DISULFID	1987	1987	BY SIMILARITY.
DISULFID	310	BY SIMILARITY.	FT	DISULFID	1988	1988	BY SIMILARITY.
DISULFID	315	BY SIMILARITY.	FT	DISULFID	1989	1989	BY SIMILARITY.
DISULFID	333	BY SIMILARITY.	FT	DISULFID	1990	1990	BY SIMILARITY.
DISULFID	352	BY SIMILARITY.	FT	DISULFID	1991	1991	BY SIMILARITY.
DISULFID	376	BY SIMILARITY.	FT	DISULFID	1992	1992	BY SIMILARITY.
DISULFID	392	BY SIMILARITY.	FT	DISULFID	1993	1993	BY SIMILARITY.
DISULFID	397	BY SIMILARITY.	FT	DISULFID	1994	1994	BY SIMILARITY.
DISULFID	414	BY SIMILARITY.	FT	DISULFID	1995	1995	BY SIMILARITY.
DISULFID	431	BY SIMILARITY.	FT	DISULFID	1996	1996	BY SIMILARITY.
DISULFID	453	BY SIMILARITY.	FT	DISULFID	1997	1997	BY SIMILARITY.
DISULFID	468	BY SIMILARITY.	FT	DISULFID	1998	1998	BY SIMILARITY.
DISULFID	473	BY SIMILARITY.	FT	DISULFID	1999	1999	BY SIMILARITY.
DISULFID	490	BY SIMILARITY.	FT	DISULFID	2000	2000	BY SIMILARITY.
DISULFID	505	BY SIMILARITY.	FT	DISULFID	2001	2001	BY SIMILARITY.
DISULFID	509	BY SIMILARITY.	FT	DISULFID	2002	2002	BY SIMILARITY.
DISULFID	522	BY SIMILARITY.	FT	DISULFID	2003	2003	BY SIMILARITY.
DISULFID	549	BY SIMILARITY.	FT	DISULFID	2004	2004	BY SIMILARITY.
DISULFID	556	BY SIMILARITY.	FT	DISULFID	2005	2005	BY SIMILARITY.
DISULFID	571	BY SIMILARITY.	FT	DISULFID	2006	2006	BY SIMILARITY.
DISULFID	586	BY SIMILARITY.	FT	DISULFID	2007	2007	BY SIMILARITY.
DISULFID	591	BY SIMILARITY.	FT	DISULFID	2008	2008	BY SIMILARITY.
DISULFID	604	BY SIMILARITY.	FT	DISULFID	2009	2009	BY SIMILARITY.
DISULFID	613	BY SIMILARITY.	FT	DISULFID	2010	2010	BY SIMILARITY.
DISULFID	618	BY SIMILARITY.	FT	DISULFID	2011	2011	BY SIMILARITY.
DISULFID	636	BY SIMILARITY.	FT	DISULFID	2012	2012	BY SIMILARITY.
DISULFID	652	BY SIMILARITY.	FT	DISULFID	2013	2013	BY SIMILARITY.
DISULFID	659	BY SIMILARITY.	FT	DISULFID	2014	2014	BY SIMILARITY.
DISULFID	675	BY SIMILARITY.	FT	DISULFID	2015	2015	BY SIMILARITY.
DISULFID	691	BY SIMILARITY.	FT	DISULFID	2016	2016	BY SIMILARITY.
DISULFID	696	BY SIMILARITY.	FT	DISULFID	2017	2017	BY SIMILARITY.
DISULFID	713	BY SIMILARITY.	FT	DISULFID	2018	2018	BY SIMILARITY.
DISULFID	729	BY SIMILARITY.	FT	DISULFID	2019	2019	BY SIMILARITY.
DISULFID	734	BY SIMILARITY.	FT	DISULFID	2020	2020	BY SIMILARITY.
DISULFID	751	BY SIMILARITY.	FT	DISULFID	2021	2021	BY SIMILARITY.
DISULFID	767	BY SIMILARITY.	FT	DISULFID	2022	2022	BY SIMILARITY.
DISULFID	772	BY SIMILARITY.	FT	DISULFID	2023	2023	BY SIMILARITY.
DISULFID	789	BY SIMILARITY.	FT	DISULFID	2024	2024	BY SIMILARITY.
DISULFID	806	BY SIMILARITY.	FT	DISULFID	2025	2025	BY SIMILARITY.
DISULFID	811	BY SIMILARITY.	FT	DISULFID	2026	2026	BY SIMILARITY.
DISULFID	828	BY SIMILARITY.	FT	DISULFID	2027	2027	BY SIMILARITY.
DISULFID	844	BY SIMILARITY.	FT	DISULFID	2028	2028	BY SIMILARITY.
DISULFID	849	BY SIMILARITY.	FT	DISULFID	2029	2029	BY SIMILARITY.
DISULFID	892	BY SIMILARITY.	FT	DISULFID	2030	2030	BY SIMILARITY.
DISULFID	908	BY SIMILARITY.	FT	DISULFID	2031	2031	BY SIMILARITY.
DISULFID	913	BY SIMILARITY.	FT	DISULFID	2032	2032	BY SIMILARITY.
DISULFID	930	BY SIMILARITY.	FT	DISULFID	2033	2033	BY SIMILARITY.
DISULFID	946	BY SIMILARITY.	FT	DISULFID	2034	2034	BY SIMILARITY.
DISULFID	952	BY SIMILARITY.	FT	DISULFID	2035	2035	BY SIMILARITY.
DISULFID	968	BY SIMILARITY.	FT	DISULFID	2036	2036	BY SIMILARITY.
DISULFID	984	BY SIMILARITY.	FT	DISULFID	2037	2037	BY SIMILARITY.
DISULFID	989	BY SIMILARITY.	FT	DISULFID	2038	2038	BY SIMILARITY.
DISULFID	1011	BY SIMILARITY.	FT	DISULFID	2039	2039	BY SIMILARITY.
DISULFID	1020	BY SIMILARITY.	FT	DISULFID	2040	2040	BY SIMILARITY.
DISULFID	1211	BY SIMILARITY.	FT	DISULFID	2041	2041	BY SIMILARITY.
DISULFID	1231	BY SIMILARITY.	FT	DISULFID	2042	2042	BY SIMILARITY.

InterPro; IPR001512; Ask_hydroxy_1_S.	
InterPro; IPR0001774; DSL.	
InterPro; IPR000742; EGF 2.	
InterPro; IPR001881; EGF_Ca.	
InterPro; IPR001438; EGF_I1.	
InterPro; IPR0006209; EGF_Ilike.	
InterPro; IPR002049; Laminin EGF.	
InterPro; IPR009041; PMP_inhibitor.	
InterPro; IPR001007; VWF_C.	
Pfam; PF01414; DSL; 1.	
Pfam; PF00008; EGF; 14.	
PRINTS; PR00010; EGFBLOOD.	
PRINTS; PR00011; EGFLAMININ.	
SMART; SM00051; DSL; 1.	
SMART; SM00179; EGF_CA; 10.	
SMART; SM00214; VWF_C; 1.	
PROSITE; PS00010; ASX_HYDROXYL; 10.	
PROSITE; PS00022; EGF_I; 16.	
PROSITE; PS00186; EGF_2; 12.	
PROSITE; PS00026; EGP_3; 15.	
PROSITE; PS01187; EGF_CA; 8.	
PROSITE; PS01208; VWF_C; 1; FALSE_NEG.	
Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein; Receptor; Transmembrane; Signal.	
KW SIGNAL 1 26	POTENTIAL.
KW CHAIN 27 1213 JAGGED 3.	
KW DOMAIN 27 1064 EXTRACELLULAR (POTENTIAL).	
KW TRANSMEM 1065 1087 CYTOSOLIC (POTENTIAL).	
KW DOMAIN 1088 1213 CYTOSMERIC (POTENTIAL).	
KW DOMAIN 164 226 DSL.	
KW DOMAIN 227 260 EGF-LIKE 1.	
KW DOMAIN 258 291 EGF-LIKE 2.	
KW DOMAIN 293 331 EGF-LIKE 3.	
KW DOMAIN 333 369 EGF-LIKE 4.	
KW DOMAIN 371 407 EGF-LIKE 5 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 409 445 EGF-LIKE 6 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 447 482 EGF-LIKE 7 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 484 520 EGF-LIKE 8 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 522 558 EGF-LIKE 9.	
KW DOMAIN 592 624 EGF-LIKE 10.	
KW DOMAIN 626 662 EGF-LIKE 11 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 664 700 EGF-LIKE 12 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 702 738 EGF-LIKE 13.	
KW DOMAIN 746 777 EGF-LIKE 14 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 779 815 EGF-LIKE 15 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 817 853 EGF-LIKE 16 CALCIUM-BINDING (POTENTIAL).	
KW DOMAIN 860 914 VWF_C.	
KW DOMAIN 918 956 EGF-LIKE 17.	
KW DISULFID 231 242 BY SIMILARITY.	
KW DISULFID 235 248 BY SIMILARITY.	
KW DISULFID 250 259 BY SIMILARITY.	
KW DISULFID 262 273 BY SIMILARITY.	
KW DISULFID 268 279 BY SIMILARITY.	
KW DISULFID 281 290 BY SIMILARITY.	
KW DISULFID 297 309 BY SIMILARITY.	
KW DISULFID 303 319 BY SIMILARITY.	
KW DISULFID 321 330 BY SIMILARITY.	
KW DISULFID 337 348 BY SIMILARITY.	
KW DISULFID 342 357 BY SIMILARITY.	
KW DISULFID 359 368 BY SIMILARITY.	
KW DISULFID 375 396 BY SIMILARITY.	
KW DISULFID 380 395 BY SIMILARITY.	
KW DISULFID 397 406 BY SIMILARITY.	
KW DISULFID 413 424 BY SIMILARITY.	
KW DISULFID 418 433 BY SIMILARITY.	
KW DISULFID 435 444 BY SIMILARITY.	
KW DISULFID 451 461 BY SIMILARITY.	
KW DISULFID 455 470 BY SIMILARITY.	
KW DISULFID 472 481 BY SIMILARITY.	
KW DISULFID 488 499 BY SIMILARITY.	
KW DISULFID 493 508 BY SIMILARITY.	
KW DISULFID 510 519 BY SIMILARITY.	
KW DISULFID 526 537 BY SIMILARITY.	
FT DISULFID 531 546 BY SIMILARITY.	
FT DISULFID 548 557 BY SIMILARITY.	
FT DISULFID 630 641 BY SIMILARITY.	
FT DISULFID 635 650 BY SIMILARITY.	
FT DISULFID 652 661 BY SIMILARITY.	
FT DISULFID 668 679 BY SIMILARITY.	
FT DISULFID 673 688 BY SIMILARITY.	
FT DISULFID 690 699 BY SIMILARITY.	
FT DISULFID 706 717 BY SIMILARITY.	
FT DISULFID 711 726 BY SIMILARITY.	
FT DISULFID 728 737 BY SIMILARITY.	
FT DISULFID 734 746 BY SIMILARITY.	
FT DISULFID 742 753 BY SIMILARITY.	
FT DISULFID 750 763 BY SIMILARITY.	
FT DISULFID 769 783 BY SIMILARITY.	
FT DISULFID 788 803 BY SIMILARITY.	
FT DISULFID 805 814 BY SIMILARITY.	
FT DISULFID 821 832 BY SIMILARITY.	
FT DISULFID 826 841 BY SIMILARITY.	
FT DISULFID 843 852 BY SIMILARITY.	
FT DOMAIN 938 941 POLY-PRO.	
FT CARBOYD 139 139 N-LINKED (GLCNAC . . ) (POTENTIAL).	
FT CARBOYD 214 214 N-LINKED (GLCNAC . . ) (POTENTIAL).	
FT CARBOYD 556 556 N-LINKED (GLCNAC . . ) (POTENTIAL).	
FT CARBOYD 742 742 N-LINKED (GLCNAC . . ) (POTENTIAL).	
FT CARBOYD 957 957 N-LINKED (GLCNAC . . ) (POTENTIAL).	
FT CARBOYD 988 988 N-LINKED (GLCNAC . . ) (POTENTIAL).	
FT CARBOYD 1042 1042 N-LINKED (GLCNAC . . ) (POTENTIAL).	
FT SEQUENCE 1213 AA; 1213 3365 MW; 5CF16A7B20D9534 CRC64; SQ SEQUENCE 1213 AA; 1213 3365 MW; 5CF16A7B20D9534 CRC64;	
Query Match 12.1%; Score 91.5; DB 1; Length 1213; Best Local Similarity 24.8%; Pred. No. 0.23; Mismatches 50; Indels 37; Gaps 8; Matches 35; Conservative 35;	
Qy 17 ACAMLLC-HGSLQHTFOOHHLRP--EGGTGBV---IAAHRCCKNKRIBERSQTYKC 67	
Db 336 ACUSNPANGGTTKETSGTGGTCATGWSGTSEINVDDCTPNQCKHGTCDLVNGFKC 395	
Qy 68 SCLP--GKV---AGTRNRSPCVDAISI--VIGKWNCEMEP-----100	
Db 396 ACPHWTKTCKQIDANZCSDKP-CVNAKSCHNLIGAYFCCLPGWSQNCDINNDCKGQ 454	
RESULT 7	
ID DP87_DICDI DP87_DICDI STANDARD PRT; 555 AA.	
AC Q04503; 01-FEB-1994 (Rel. 28, Created)	
DT 01-FEB-1994 (Rel. 28, Last sequence update)	
DT 15-MAR-2004 (Rel. 43, Last annotation update)	
DE Prespore DP87 precursor.	
GN COND OR DP87.	
OS Dictyostelium discoideum (Slime mold).	
OC Eukaryota: Mycetozoa; Dictyosteliida; Dictyostelium.	
OX NCBI_TaxID:44689; [1]	
RN SEQUENCE FROM N_A.	
RC STRAIN=AX3;	
RX MEDLINE:94008533; PubMed=8404532;	
RA Ozaki T., Nakao H., Orii H., Morio T., Takeuchi I., Tasaka M.,	
RT "Developmental regulation of transcription of a novel prespore-specific gene (dp87) in Dictyostelium discoideum.";	
RL Development 117:1299-1308 (1993).	
CC -!- SUBCELLULAR LOCATION: Stored in prespore vacuoles until it is discharged into the interspacers of spores during spore formation.	
CC -!- INDUCTION: By exogenous cAMP, repressed by DIF.	
CC -!- SIMILARITY: Contains 7 prespore motif repeats.	
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way	

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

- CC STRAIN=Berkeley; PubMed=20196016; ID NOTC\_DROME\_STANDARD; PRT; 2703 AA.  
 CC ADAMS M.D., Celinkin K.E., Evans C.A., Gocayne J.D.,  
 CC FISHBACK P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 CC Amanatides P.G., Scherer S.E., Richards S.D., Henderson S.N.,  
 CC GEORGE R.A., Lewis S.E., Richards S.D., Ashburner M., Henderson S.N.,  
 CC SUTTON G.G., WORTMAN J.R., YANDELL M.D., ZHANG Q., CHEN L.X.,  
 CC BRANDON R.C., ROGERS Y.-H.C., BLAZEY R.G., CHAMPE M., PEIFFER B.D.,  
 CC WAN K.H., DOYLE C., BAXTER E.G., HILT G., NELSON C.R., MIKLOS G.I.G.,  
 CC ALEXANDRE P.G., BEASLEY E.M., BALDWIN D.,  
 CC DRAGICEVIC C., ANDREW A.H.-J., ANDREW A.H.-J.,  
 CC BALLEW R.M., BASU A., BAXENDALE J., BHANDARI D., BOILSHAKOV S.,  
 CC BEESON K.Y., BENOS P.V., BERMAN B.P., BROCKLEBURN P., BROTTIER P.,  
 CC BORKOVKA D., BOTCHAN M.R., BOUCH L.,  
 CC BURTIS K.C., BUSAM D.A., BUTLER H., CADIEU E., CENTER A., CHANDRA I.,  
 CC CHERRY J.M., COWLEY S., DAHLKE C., DAVENPORT L., DAWES P.,  
 CC DE PABLOS B., DELICHER A., DENG Z., DIAZ S.M., DIAZ S.M.,  
 CC DODSON K., DOUP L.E., DOWNEY M., DUGAN-ROCHA S., DUNKOV B.C., DUNN P.,  
 CC DURBIN K.J., EVANGELISTI C.C., FERRAZ C., FERRIERA S., FLEISCHMANN W.,  
 CC FOSLER C., GABRIELINI A.E., GARG N.S., GELBART W.M., GLASSER K.,  
 CC GLODEK A., GONG F., GORRELL J.H., GU Z., GUAN P., HARRIS M.,  
 CC HARRIS N.L., HARVEY D.A., HEIMAN T.J., HERNANDEZ J.R., HOUCY J.,  
 CC HOSTIN D., HOUSTON K.A., HOWLAND T.J., WEI M.-H., IBEGMAM C.,  
 CC JALALI M., KALUSH F., KARPEN G.H., KE Z., KENNISON J.A., KETCHUM K.A.,  
 CC KIMMEL B.E., KOIDE R.C.D., KRAFT C., KRAVITZ S., KULP D., LAI Z.,  
 CC LASKO P., LEI Y., LEVITSKY A.A., LI J.H., LI Z., LIANG Y., LIN X.,  
 CC LIU X., MATTEI B., MCINTOSH T.C., MCLEOD M.P., MCLEOD M.P., MCHARRON D.,  
 CC MERKULOV G., MILLASHINA N.V., MOBARRY C., MORRIS J., MOSIREFI A.,  
 CC MOUNT S.M., MOY M., MURPHY L., MURPHY D.M., NELSON D.L.,  
 CC NELSON D.R., NELSON K.A., NIXON K., NUSSBERGER D.R., PACIENZI J.M.,  
 CC PALAZZOLO M., PITTMAN G.S., PAN S., POLLARD J., PURI V., REESE M.G.,  
 CC REINEST K., REMMINGHORN K., SAUNDERS R.D.C., SCHHEELER F., SHEN H.,  
 CC SHUE B.C., SIDEN-KIANOS I., SIMPSON M., SKUPSKI M.P., SMITH T.,  
 CC SPIER E., SPRADING A.C., STAPLETON M., STRONG R., SUN E.,  
 CC SWIRSKA R., TECTOR C., TURNER R., WENTER E., WANG A.H., WANG X.,  
 CC WANG Z.-Y., WASBARMAN D.A., WEINSTOCK G.M., WEISSENBACH J.,  
 CC YE J., YE H.-F., WOODAGE T., WU D., YANG S., YAO Q.-A.,  
 CC ZHENG X.H., ZHONG F.N., ZHAO M., ZHANG G., ZHAO Q., ZHENG L.,  
 CC GIBBS R.A., MYERS E.W., RUBIN G.M., VENTER J.C.;  
 CC "The genome sequence of *Drosophila melanogaster*,"  
 CC Science 287:2185-2195 (2000).  
 CC [4]
- Qy SEQUENCE FROM N.A.  
 Qy STRAIN=Oregon\_R;  
 Qy MEDLINE=20196017; ID NOTC\_DROME\_STANDARD; PRT; 2703 AA.  
 Db 238 HHLHR---PEGGTCEVIAAHRCNNKRTERSQTV-----KCSCLPGKVAGTRNRSPC 84  
 Db 238 HNLCRGFQCPGSHCEVLEKHPVCVRNIVPHPPPPPOICGSYNCGPYIICTINGHPTC 297  
 Qy 85 VDASIVIGKWWCMEPCLEGEECKTLP-----DNSGN-MCATGNKNIKUTRTHP 131  
 Db 298 IRGD---GYLCNQTRCPHDYQCETISTNVKGSKPKNDECKWHRCPGSSCFNSRNGP 351  
 Qy SEQUENCE FROM N.A.  
 Qy STRAIN=Oregon\_R; TISSUE=Embryo;  
 Rx MEDLINE=86709539; ID NOTC\_DROME\_STANDARD; PRT; 2703 AA.  
 Rx WHARTON K.A., JOHANSEN K.M., XU T., ARTAVANIS-TSAKONAS S.;  
 Rx "Nucleotide sequence from the neurogenic locus notch implies a gene product that shares homology with proteins containing EGR-like repeats,"  
 Rx Cell 43:567-581 (1985).  
 Rx SEQUENCE FROM N.A.  
 Rx STRAIN=Canton-S, and Oregon-R; TISSUE=Embryo;  
 Rx MEDLINE=87064624; ID NOTC\_DROME\_STANDARD; PRT; 2703 AA.  
 Rx KIDD S., KELLEY M.R., YOUNG M.W., "Sequence of the notch locus of *Drosophila melanogaster*," Mol. Cell. Biol. 6:3094-3108 (1986).  
 Rx SEQUENCE FROM N.A.
- CC STRAIN=Berkeley; PubMed=10731132;  
 CC MEDLINE=20196016; ID NOTC\_DROME\_STANDARD; PRT; 2703 AA.  
 CC ADRIEL J.F., AGHAYANI A., BAXENDALE J., BHANDARI D., BOILSHAKOV S.,  
 CC BALLEW R.M., BASU A., BAXENDALE J., BHANDARI D., BOILSHAKOV S.,  
 CC BEESON K.Y., BENOS P.V., BERMAN B.P., BROCKLEBURN P., BROTTIER P.,  
 CC BORKOVKA D., BOTCHAN M.R., BOUCH L.,  
 CC BURTIS K.C., BUSAM D.A., BUTLER H., CADIEU E., CENTER A., CHANDRA I.,  
 CC CHERRY J.M., COWLEY S., DAHLKE C., DAVENPORT L., DAWES P.,  
 CC DE PABLOS B., DELICHER A., DENG Z., DIAZ S.M., DIAZ S.M.,  
 CC DODSON K., DOUP L.E., DOWNEY M., DUGAN-ROCHA S., DUNKOV B.C., DUNN P.,  
 CC DURBIN K.J., EVANGELISTI C.C., FERRAZ C., FERRIERA S., FLEISCHMANN W.,  
 CC FOSLER C., GABRIELINI A.E., GARG N.S., GELBART W.M., GLASSER K.,  
 CC GLODEK A., GONG F., GORRELL J.H., GU Z., GUAN P., HARRIS M.,  
 CC HARRIS N.L., HARVEY D.A., HEIMAN T.J., HERNANDEZ J.R., HOUCY J.,  
 CC HOSTIN D., HOUSTON K.A., HOWLAND T.J., WEI M.-H., IBEGMAM C.,  
 CC JALALI M., KALUSH F., KARPEN G.H., KE Z., KENNISON J.A., KETCHUM K.A.,  
 CC KIMMEL B.E., KOIDE R.C.D., KRAFT C., KRAVITZ S., KULP D., LAI Z.,  
 CC LASKO P., LEI Y., LEVITSKY A.A., LI J.H., LI Z., LIANG Y., LIN X.,  
 CC LIU X., MATTEI B., MCINTOSH T.C., MCLEOD M.P., MCLEOD M.P., MCHARRON D.,  
 CC MERKULOV G., MILLASHINA N.V., MOBARRY C., MORRIS J., MOSIREFI A.,  
 CC MOUNT S.M., MOY M., MURPHY L., MURPHY D.M., NELSON D.L.,  
 CC NELSON D.R., NELSON K.A., NIXON K., NUSSBERGER D.R., PACIENZI J.M.,  
 CC PALAZZOLO M., PITTMAN G.S., PAN S., POLLARD J., PURI V., REESE M.G.,  
 CC REINEST K., REMMINGHORN K., SAUNDERS R.D.C., SCHHEELER F., SHEN H.,  
 CC SHUE B.C., SIDEN-KIANOS I., SIMPSON M., SKUPSKI M.P., SMITH T.,  
 CC SPIER E., SPRADING A.C., STAPLETON M., STRONG R., SUN E.,  
 CC SWIRSKA R., TECTOR C., TURNER R., WENTER E., WANG A.H., WANG X.,  
 CC WANG Z.-Y., WASBARMAN D.A., WEINSTOCK G.M., WEISSENBACH J.,  
 CC YE J., YE H.-F., WOODAGE T., WU D., YANG S., YAO Q.-A.,  
 CC ZHENG X.H., ZHONG F.N., ZHAO M., ZHANG G., ZHAO Q., ZHENG L.,  
 CC GIBBS R.A., MYERS E.W., RUBIN G.M., VENTER J.C.;  
 CC "The genome sequence of *Drosophila melanogaster*,"  
 CC Science 287:2220-2222 (2000).  
 CC [5]
- CC SEQUENCE OF 1-8 FROM N.A.  
 CC MEDLINE=85099322; ID NOTC\_DROME\_STANDARD; PRT; 2703 AA.  
 CC WHARTON K.A., YEDVOBNICK B., FINNERY V.G., ARTAVANIS-TSAKONAS S.;  
 CC "Opal: a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in *D. melanogaster*,"  
 CC Cell 40:55-62 (1985).  
 CC [6]
- CC SEQUENCE OF 1-8 FROM N.A.  
 CC MEDLINE=87257846; ID NOTC\_DROME\_STANDARD; PRT; 2703 AA.  
 CC KELLEY M.R., KIDD S., BERTI R.L., YOUNG M.W., "Restriction of P-element insertions at the Notch locus of *Drosophila melanogaster*," Mol. Cell. Biol. 7:1545-1548 (1987).  
 CC [7]
- CC INTERACTION WITH DX, AND MUTANT SU42C.  
 CC RP

REVIEW  
Diederich R.J., Matsuno K., Hing H., Artavanis-Tsakonas S.;  
"Cytosolic interaction between delta<sub>x</sub> and Notch ankyrin repeats  
implicates delta<sub>x</sub> in the Notch signaling pathway.",  
Development 120:473-481(1994). [8]

INTERACTION WITH DX.  
MEDLINE=95401878; PubMed=7671825;  
Matsuno K., Diederich R.J., Go M.J., Blaumueller C.M.,  
Artavanis-Tsakonas S.;  
"Delta<sub>x</sub> acts as a positive regulator of Notch signaling through  
interactions with the Notch ankyrin repeats.",  
Development 121:2633-2644(1995). [9]

S3 CLEAVAGE BY PSN.  
MEDLINE=99221487; PubMed=10206646;  
Struhrl G., Greenwald I.;  
"Presenilin is required for activity and nuclear access of Notch in  
Drosophila.",  
Nature 398:522-525(1999). [10]

S3 CLEAVAGE BY PSN.  
MEDLINE=99221488; PubMed=10206647;  
Ye Y., Lukinova N., Portini M.E.;  
"Neurogenetic phenotypes and altered Notch processing in Drosophila  
mutants.",  
Nature 398:525-529(1999). [11]

S2 CLEAVAGE BY KUZ.  
MEDLINE=21657146; PubMed=11799064;  
Lieber T., Kidd S., Young M.W.;  
"Kuzbanian-mediated cleavage of Drosophila Notch.",  
Genes Dev. 16:209-221(2002). [12]

MUTANT MCDS  
MEDLINE=21575956; PubMed=11719244;  
Romain P., Khechumian K., Seugnet L., Arbogast N., Ackermann C.,  
Heitzler P.;  
"Novel Notch alleles reveal a Delta<sub>x</sub>-dependent pathway repressing  
neural fate.",  
Curr. Biol. 11:1729-1738(2001). [13]

REVIEW.  
MEDLINE=222256570; PubMed=12369105;  
Portin P.;  
"General outlines of the molecular genetics of the Notch signalling  
pathway in Drosophila melanogaster: a review.",  
Hereditas 116:89-96(2002).

-!- FUNCTION: Signaling protein, which regulates, with both positive  
and negative signals the differentiation of, at least central and  
peripheral nervous system and eye, wing disk, oogenesis, segment  
appendages such as antennae and legs, and muscles, through laterally  
inhibition or induction. Functions as a receptor for membrane-  
bound ligands Delta and Serrate to regulate cell-fate  
determination. Upon ligand activation, and releasing from the ce-  
membrane, the Notch intracellular domain (NICD) forms a  
transcriptional activator complex with Su(H) (Suppressor of  
hairless) and activates genes of the E(spl) complex. Essential for  
proper differentiation of ectoderm.

-!- SUBUNIT: Interacts with Su(H) when activated. Interacts with Dx  
via its ANK repeats.

-!- SUBCELLULAR LOCATION: Type I membrane protein. Upon activation at  
S3 cleavage, it is released from the cell membrane and enters in  
the nucleus in conjunction with Su(H).

-!- PTM: Upon binding its ligands such as Delta or Serrate, it is  
cleaved (S3 cleavage) in its extracellular domain, close to the  
transmembrane domain. S2 cleavage is probably mediated by Kuz. I  
is then cleaved (S3 cleavage) downstream of its transmembrane  
domain, releasing it from the cell membrane. S3 cleavage require  
Psn.

-!- SIMILARITY: Belongs to the NOTCH family.

-!- SIMILARITY: Contains 36 EGF-like domains.

-!- SIMILARITY: Contains 36 EGF-like domains.

-> SIMILARITY: Contains 6 ANK repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).

CC -----

DR AAB59220; 1;  
 DR EMBL; M16152; AAB59220\_1; JOINED.  
 DR EMBL; M16149; AAB59220\_1; JOINED.  
 DR EMBL; M16150; AAB59220\_1; JOINED.  
 DR EMBL; M16151; AAB59220\_1; JOINED.  
 DR EMBL; K03508; AAB28725\_1; -.  
 DR EMBL; M13689; AAB28725\_1; JOINED.  
 DR EMBL; K03507; AAB28725\_1; JOINED.  
 DR EMBL; AB003426; AAF45848\_2.; -.  
 DR EMBL; AL035436; CAB37610\_1; -.  
 DR EMBL; AL035435; CAB37610\_1; JOINED.  
 DR EMBL; M12175; AAF74496\_1; -.  
 DR EMBL; M16025; AAA28726\_1; -.

Query Match 11.5%; Score 86.5; DB 1; Length 2703;  
 Best Local Similarity 25.9%; Pred. No. 1.7; Mismatches 30; Indels 19; Gaps 3;  
 Matches 22; Conservative 10; MisMatches 34;

Qy 40 EGGTCEV---IAAHRCNNRNLIEERTSQTVKESCLPGKVA-----GTTNRPS 83  
 Db 1214 QGQNCELNIDDCAPNP|CONGGTCHDYRMNFSCSPPSTMGLICEINKKDCRPGACTINGS 1273

Qy 84 CVDASIVIGKWWCEMPEPLEGECCK 108  
 Db 1274 CIDR---VGGFEBCVCPQPGFVGARCE 1295

RESULT 9

DLI4	HUMAN	STANDARD	PRT	685 AA.
ID	DLI4_HUMAN			
AC	Q9NR61; Q9NQT9;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Delta-like protein 4 precursor (Drosophila Delta homolog 4)			
DE	(UNCG1895/PRO34341).			
GN	DLI4.			
OS	Home sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarhini; Hominoidea; Homo.			
OX	NCB_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=0296719; Published=1083704;			
RA	Shutter J.R., Scully S., Fan W., Richards W.G., Kitajewski J.,			
RA	Deblaudre G.A., Kintner C.R., Stark K.L.;			
RT	"DLI4, a novel Notch ligand expressed in arterial endothelium.",			
RL	Genes Dev. 14:1313-1318(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Sakano S.;			
RA	Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.			

Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Hellens S., Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J., Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A., Vandlen R., Watarabe C., Wileand D., Woods M.-H., Yamura D., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goodard A., Wood W.I., Godowski P.; "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment"; *Biochem Res. Lett.* 13:2265-2270 (2003). [5]

SEQUENCE OF 33-665 FROM N.A.  
RC TISSUE=Placenta;  
RA Mailhos C., Modlich U., Harris J., Lewis J., Bicknell R., RA Ish-Horowicz D.;  
RT "A novel Delta gene expressed in embryonic and tumour vasculature.";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Plays a role in the Notch signaling pathway. Activates  
CC Notch-1 and Notch-4 (By similarity).  
CC -!- SUBUNIT: Binds to Notch-1 and Notch-4 (By Similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).  
CC -!- TISSUE SPECIFICITY: Expressed in vascular endothelium.  
CC -!- DOMAIN: The Delta-Serrate-Lag2 (DSL) domain is required for  
CC binding to the Notch receptor.  
CC -!- SIMILARITY: Contains 8 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 DSL domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR AF253468; AAP76427.1; -  
EMBL; AB036331; BAB16085.1; -  
DR AB04894; BAB16881.1; -  
EMBL; AY358894; AAQ89253.1; -  
EMBL; AF279405; AAP791912.1; -  
PIR; JC7570; JC7570.  
DR HSSP; P00740; 1EDM.  
Genew; HGNC:2910; DLL4.  
MIM; 605155; -  
DR GO; GO:0005112; F:Notch binding; TAS.  
DR GO; GO:0008015; P:circulation; TAS.  
DR GO; GO:007165; P:sign transduction; TAS.  
DR InterPro; IPR000152; Asx\_hydroxyl\_S.  
DR InterPro; IPR001774; DSL.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR01438; EGF\_II.  
DR InterPro; IPR006209; EGF\_Like.  
DR Pfam; PRO114; DSL; 1.  
DR Pfam; PF00088; EGF\_7.  
DR PRINTS; PRO0010; EGFBLOOD.  
DR SMART; SM00051; DSL; 1.  
DR SMART; SM00179; EGF\_CA; 2.  
DR PROSITE; PS000010; ASX\_HYDROXYL; 1.  
DR PROSITE; PS00022; EGF\_1; 8.  
DR PROSITE; PS01186; EGF\_2; 7.  
KW Differentiation; Glycoprotein.  
FT SIGNAL\_1; 26 POTENTIAL.  
FT CHAIN; 27 DELTA-LIKE PROTEIN 4.  
FT DOMAIN; 27 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM; 530 POTENTIAL.  
FT DOMAIN; 551 CYTOPLASMIC (POTENTIAL).  
-----  
DR DOMAIN; 155 DSL.  
FT DOMAIN; 217 EGF-LIKE 1.  
FT DOMAIN; 251 EGF-LIKE 2.  
FT DOMAIN; 282 EGF-LIKE 3.  
FT DOMAIN; 322 EGF-LIKE 4.  
FT DOMAIN; 360 EGF-LIKE 5.  
FT DOMAIN; 400 EGF-LIKE 6.  
FT DOMAIN; 438 EGF-LIKE 7.  
FT DOMAIN; 440 EGF-LIKE 8.  
FT DOMAIN; 480 518 BY SIMILARITY.  
FT DISULFID; 222 233 BY SIMILARITY.  
FT DISULFID; 226 239 BY SIMILARITY.  
FT DISULFID; 241 250 BY SIMILARITY.  
FT DISULFID; 253 264 BY SIMILARITY.  
FT DISULFID; 259 270 BY SIMILARITY.  
FT DISULFID; 272 281 BY SIMILARITY.  
FT DISULFID; 288 300 BY SIMILARITY.  
FT DISULFID; 294 310 BY SIMILARITY.  
FT DISULFID; 312 321 BY SIMILARITY.  
FT DISULFID; 328 339 BY SIMILARITY.  
FT DISULFID; 333 348 BY SIMILARITY.  
FT DISULFID; 350 359 BY SIMILARITY.  
FT DISULFID; 366 377 BY SIMILARITY.  
FT DISULFID; 371 388 BY SIMILARITY.  
FT DISULFID; 390 399 BY SIMILARITY.  
FT DISULFID; 406 417 BY SIMILARITY.  
FT DISULFID; 411 426 BY SIMILARITY.  
FT DISULFID; 428 437 BY SIMILARITY.  
FT DISULFID; 444 464 BY SIMILARITY.  
FT DISULFID; 466 475 BY SIMILARITY.  
FT DISULFID; 484 495 BY SIMILARITY.  
FT DISULFID; 489 506 BY SIMILARITY.  
FT CARBOYD; 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOYD; 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOYD; 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOYD; 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 685 AA; 74604 MW; 6CF89D3C220AAC89 CRC64;

Query Match 11.4%; Score 86; DB: 1; Length 685;  
Best Local Similarity 25.0%; Pred. No. 0.45; Mismatches 9; Indels 30; Gaps 6;  
Matches 28; Conservative 28; Score 100.0%.

Qy 14 WISAC---AMILCHGSLOHTFOQHHLHRPQQGTCVIAHRCNCNKRIBERSQTVKCSC 69  
DB 268 WQCTCDEGWGLFCDDDLNYC--THESPKRNATC-----SNSGSRSTC-CRC 312

Qy 70 LPGKIVAGTTNRPSVCDATVIGRWCEMPCLEGECCKTLPDNSGWMCATG 121  
DB 313 RPYGTG-----VDCELELSE--CDSNPCRNGGSKDQEDGYHCLCPG 353

RESULT 10  
JAG1\_HUMAN  
ID JAG1\_HUMAN STANDARD; PRT; 1218 AA.  
AC P78504; O14902; O15122; Q15816;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE Jagged\_1 precursor (Jagged1) (hJ1).  
GN JAG1 OR JAGL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Butheria; Primates; Catarrhini; Hominoidea; Homo.  
RN [1] NCBI\_TaxID=9606;  
RN [1] SEQUENCE FROM N.A.  
RP MEDLINE=97422615; PubMed=9268641;  
RX Oda T.; Elkahloun A.G.; Melitter P.S.; Chandrasekharappa S.C.;  
RT "Identification and cloning of the human homolog (JAG1) of the rat  
RL Genomics 43:376-379(1997)."  
RN [2] SEQUENCE FROM N.A., AND VARIANT AGS CYS-184.

- K** TISSUE=Bone marrow; PubMed=9207788;  
 RX MEDLINE=97351506; PubMed=9207787;  
 RA Oda T., Elkahloun A.G., Pike B.I., Okajima K., Krantz I.D., Genin A.,  
 RA Piccoli D.A., Meltzer P.S., Spinner N.B., Collins F.S.,  
 RA Chandrasekharappa S.C.;  
 RA "Mutations in the human Jagged1 gene are responsible for Alagille  
 syndrome." ;  
 RT Nat. Genet. 16:235-242 (1997).  
 RL [9]  
 RN RN  
 RP RP  
 SEQUENCE FROM N.A., AND FUNCTION.  
 RX MEDLINE=9812342; PubMed=946510;  
 RA Li L., Milner L.A., Deng Y., Iwata M., Banta A.B., Graf L.,  
 RA Marcovina S., Friedman C., Trask B.J., Hood L., Torok-Storb B.,  
 RA "The human homolog of rat Jagged1 expressed by marrow stroma inhibits  
 differentiation of 32D cells through interaction with Notch1." ;  
 RT Immunity 8:43-55 (1998).  
 RL [4]  
 RN RN  
 RP RP  
 SEQUENCE FROM N.A.  
 RX TISSUE=Cervical carcinoma;  
 RA MEDLINE=9262417; PubMed=10329626;  
 RA Bash J., Zong W.-X., Banga S., Rivera A., Ballard D.W., Ron Y.,  
 RA Gelinas C.;  
 RT "Rel/NF kappaB can trigger the Notch signaling pathway by inducing the  
 expression of Jagged1, a ligand for Notch receptors." ;  
 RL EMBO J. 18:2803-2811 (1999).  
 RN RN  
 RP RP  
 SEQUENCE FROM N.A.  
 RX MEDLINE=919180765; PubMed=110079256;  
 RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,  
 RA Banks A., Leiman J., Ward D., Ish-Horowicz D., Artavanis-Tsakonas S.;  
 RT "Human ligands of the Notch receptor." ;  
 RL Am. J. Pathol. 154:785-794 (1999).  
 RN RN  
 RP RP  
 SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA DeLousas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavridis G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakley S.E., Bridgeman A.M., Brown A.J.,  
 RA Burkitt W.D., Butler A.P., Carter C., Carter S.Y., Clark S.Y., Cleee C.M.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Connor R.E., Corby N.R.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Dunn M.,  
 RA Coulson A., Coville G.J., Deadman R.T., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Knights A., Laird G.K., Johnson D.,  
 RA Lehvaaelahti M.H., Leverhauser C., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConnaughey L.J., McElroy A.A.,  
 RA Milne S.A., Mistrey D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver J., Parker A., Patel R., Pearce T.A.V., Peck A.J.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
 RA Skuce C.L., Seward C.L., Smith M.S., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmot L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20." ;  
 RL Nature 414:865-871 (2001).  
 RN RN  
 RP RP  
 SEQUENCE OF 14-1227 FROM N.A.  
 RC TISSUE=Umbilical vein endothelial cells;  
 RX MEDLINE=97115768; PubMed=8955070;  
 RA RA Maciaig T., Pepper M.S., McMahon G.A., Nguyen F., Montesano R.,  
 RA Rogers J.;  
 RT "An antisense oligonucleotide to the notch ligand jagged enhances  
 fibroblast growth factor-induced angiogenesis in vitro." ;  
 RL J. Biol. Chem. 271:32499-32502 (1996).  
 RN RN  
 DISEASE.

CC decisions during hematopoiesis. Seems to be involved in early CC and late stages of mammalian cardiovascular development. Inhibits CC myoblast differentiation (in vitro). Enhances fibroblast growth factor-induced angiogenesis (in vitro).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: Widely expressed in adult and fetal tissues.

CC In cervix epithelium expressed in undifferentiated subcolumnar reserve cells and squamous metaplasia. Expression is up-regulated in cervical squamous cell carcinoma. Expressed in bone marrow cell-line HS-274 which supports the long-term maintenance of immature progenitor cells.

CC -!- DEVELOPMENTAL STAGE: Expressed in 32-52 days embryos in the distal cardiac outflow tract and pulmonary artery, major arteries, portal vein, optic vesicle, oesophagus, bronchial arches, metanephros, pancreas, mesocardium, around the major bronchial branches, and in the neural tube.

CC -!- DISEASE: Defects in JAG1 are the cause of Alagille syndrome (AGS) (MIM:1138450). AGS is an autosomal dominant development disorder that affects structures in the liver, heart, skeleton, eye, kidney and other organs.

CC -!- DISEASE: Defects in JAG1 are a cause of tetralogy of Fallot (TOF) (MIM:181500). TOF is a congenital heart anomaly which consists of pulmonary stenosis, ventricular septal defect, dextroposition of the aorta (aorta is on the right side instead of the left), and hypertrophy of the right ventricle. This condition results in a blue baby at birth due to inadequate oxygenation.

CC Surgical correction is emergent.

CC -!- SIMILARITY: Contains 15 EGF-like domains.

CC -!- SIMILARITY: Contains 1 DSL domain.

CC -!- CAUTION: Ref. 7 sequence differs from that shown due to a frameshift in position 1187.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between Local Similarity 23.8%; Pred. No. 0.81; Matches 34; Conservative 11; Mismatches 34; Indels 64; Gaps 9; Score 86; DB 1; Length 1218;

Query Match 37 HPGKVAGTTRN-----PSCVDASTIVG-----KWW-----CEMEP 100

Db 363 SPGWGPTCSTNTIDDCSPNCSHGGTCQD----LYNGFKCVCPPWTGKTCQLDANECAKP 420

Qy 101 CLEGECKTLPDN-----SGW 117

Db 421 CVNAKSCKNLIASYCDCLPQNM 443

Qy 37 HPGKVAGTTRN-----PSCVDASTIVG-----KWW-----CEMEP 100

Db 363 SPGWGPTCSTNTIDDCSPNCSHGGTCQD----LYNGFKCVCPPWTGKTCQLDANECAKP 420

Qy 101 CLEGECKTLPDN-----SGW 117

Db 421 CVNAKSCKNLIASYCDCLPQNM 443

RESULT 11  
JAG1\_MOUSE  
ID JAG1\_MOUSE  
AC Q90XX0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DI 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Jagged 1 precursor (Jagged1).  
GN JAG1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus. NCBI\_TaxID=10090;

OX RN [1] \_TaxID=10090;  
RN SEQUENCE FROM N.A.; AND RECEPTOR INTERACTION.  
RC STRAIN=Swiss Webster / NIH;  
RX MEDLINE=20020271; PubMed=10551863;  
RA Shirizuka K., Chiba S., Kumano K., Hosoya N., Takahashi T., Kanda Y., Hamada Y., Yaraki Y., Hirai H.;  
RT "Mouse Jagged1 physically interacts with Notch2 and other Notch receptors: assessment by quantitative methods.";

RL J. Biol. Chem. 274:32961-32969 (1999).  
RN [2] SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shearer C.M., Schuler G.D., Altzuler S.F., Zeeberg B.R., Bustard K.H., Max S.I., Wang J., Hsieh P., Hopkins R.F., Jordan H., Moore T., Farmer A.A., Rubin G.M., Hong L., Diatchenko L., Marusina K., Casavant T.L., Scheetz T.E., Stapleton M., Soares M.B., Bonaldo M.F., Casavinti P., Prange C., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnaraine P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villacon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fehay J.J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J.W., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., RA Schnetzer A., Schein J.E., Jones S.J.M., Marra M.A.; RT Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.",  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN TISSUE SPECIFICITY.  
RX MEDLINE=20025757; PubMed=10556292;  
RA Loomes K.M., Underkoffler L.A., Morabito J., Gottlieb S., Piccoli D.A., Spinner N.B., Baldwin H.S., Oakey R.J.; RT "The expression of Jagged1 in the developing mammalian heart correlates with cardiovascular disease in Alagille syndrome.",  
RL Hum. Mol. Genet. 8:2443-2449 (1999).  
CC -!- FUNCTION: Ligand for multiple Notch receptors and involved in the mediation of Notch signaling. May be involved in cell-fate decisions during hematopoiesis. Seems to be involved in early and late stages of mammalian cardiovascular development. Inhibits myoblast differentiation (By similarity). May regulate fibroblast growth factor-induced angiogenesis.  
CC -!- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Widely expressed in many tissues, with highest expression in brain, heart, muscle and thymus.  
CC -!- DEVELOPMENTAL STAGE: At 8.75-9.75 dpc expression was detected in structures that include those destined to contribute to the cardiovascular system of the adult heart. Expression was also directed in the mesencephalon and rhombencephalon.  
CC -!- DOMAIN: The DSL domain is indispensable and sufficient for binding to NOTCH2.  
CC -!- SIMILARITY: Contains 15 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 DSL domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF17109; AAH58675.1; -.  
DR EMBL; BC058675; AAH58675.1; -.  
DR HSSP; P00740; 1EDM.  
DR MGD; MGI:105416; Jag1.  
DR GO; GO:0005516; C:extracellular; ISS.  
CC GO; GO:0005509; F:calcium ion binding; NAS.  
CC GO; GO:0008083; F:growth factor activity; ISS.  
CC GO; GO:0005312; F:Notch binding; IPF.  
CC GO; GO:0005198; F:structural molecule activity; ISS.  
CC GO; GO:0001125; F:angiogenesis; ISS.  
CC GO; GO:0001079; F:cell fate determination; ISS.  
CC GO; GO:0045446; F:endothelial cell differentiation; ISS.  
CC GO; GO:0030097; F:hemopoiesis; ISS.

FT	DISURFID	655	664	BY SIMILARITY.
PT	DISURFID	671	682	BY SIMILARITY.
PT	DISURFID	676	691	BY SIMILARITY.
PT	DISURFID	693	702	BY SIMILARITY.
PT	DISURFID	709	720	BY SIMILARITY.
PT	DISURFID	714	729	BY SIMILARITY.
FT	DISURFID	731	740	BY SIMILARITY.
FT	DISURFID	748	759	BY SIMILARITY.
FT	DISURFID	753	768	BY SIMILARITY.
FT	DISURFID	770	779	BY SIMILARITY.
FT	DISURFID	784	797	BY SIMILARITY.
FT	DISURFID	791	806	BY SIMILARITY.
FT	DISURFID	808	817	BY SIMILARITY.
FT	DISURFID	824	835	BY SIMILARITY.
FT	DISURFID	846	844	BY SIMILARITY.
FT	DISURFID	847	855	BY SIMILARITY.
FT	CARBODYD	143	143	N-LINED (GLCNAC. . .) (POTENTIAL).
FT	CARBODYD	217	217	N-LINED (GLCNAC. . .) (POTENTIAL).
FT	CARBODYD	382	382	N-LINED (GLCNAC. . .) (POTENTIAL).
FT	CARBODYD	559	559	N-LINED (GLCNAC. . .) (POTENTIAL).
FT	CARBODYD	745	745	N-LINED (GLCNAC. . .) (POTENTIAL).
FT	CARBODYD	960	960	N-LINED (GLCNAC. . .) (POTENTIAL).
FT	CARBODYD	991	991	N-LINED (GLCNAC. . .) (POTENTIAL).
FT	CARBODYD	1045	1045	N-LINED (GLCNAC. . .) (POTENTIAL).
FT	CARBODYD	1064	1064	N-LINED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1218 AA;	134163 MW;	77739F892BBT93C CRC64;
Qy	Best Local Similarity	11.4%;	Score 86;	DB 1; Length 1218;
Matches	33; Conservative	27.0%;	Pred. No. 0.81;	
	Matches	33;	9; Mismatches	34; Indels 46; Gaps 9;
Qy	39	PE---CGTCIEVIAAHRC---C-NKQRRIEERSQTVKCSCLPGKVASTTRNR-----	81	
Db	325	PEGYSGPNCE-IAEHACLSDPCHNRGSKETSSGGFBECSPWPWTGPICSTINLDDCSPNNC	3.83	
Qy	82	--PSCVDASITVG----KWW-----CEMPCLEGEECKTLPDN-----SG 115		
Db	384	SHGGTCQD--LVNGFEKCCVCPQWNTGTCQLDANECAKPCYNARSCKNLIASYYCDCCLPG	4.41	
Qy	116	WM 117		
Db	442	WM 443		
RESULT 12				
JAG1_RAT	JAG1_RAT	STANDARD;	PRT;	1219 AA.
AC	Q63722;	P70640;		
DT	28-FEB-2003	(Rel. 41,	Created)	
DT	28-FEB-2003	(Rel. 41,	Last sequence update)	
DT	28-FEB-2003	(Rel. 41,	Last annotation update)	
DB	Jagged 1 precursor (Jagged1).			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniota; Euteleostomi; Rattus.			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
NCBI_TaxId=1016;				
RN	SEQUENCE FROM N.A.			
RP	PTT			
RC	TISSUE=Sciatic nerve;			
RX	MBID=NB-95211812; PubMed=7697721;			
RA	Lindahl C.B., Shawber C.J., Boulter J., Weinmaster G.;			
RT	"Jagged: a mammalian ligand that activates Notch1.";			
RL	Cell 80:90-91 (1995).			
CC	-1- FUNCTION: Ligand for multiple Notch receptors and involved in the mediation of Notch signaling. May be involved in cell-fate decisions during hematopoiesis. Enhances fibroblast growth factor-induced angiogenesis (in vitro). Seems to be involved in early and late stages of mammalian cardiovascular development.			
CC	-1- Inhibits myoblast differentiation. May regulate fibroblast growth factor-induced angiogenesis.			
CC	-1- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3			(By similarity).
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- TISSUE/SPECIFICITY: Widely expressed in a variety of tissues.			

-!- DEVELOPMENTAL STAGE: Expression is seen in E11.5-E14.5 embryos in four distinct regions of the ventricular zone in the developing spinal cord.

-!- SIMILARITY: Contains 15 EGF-like domains.

-!- SIMILARITY: Contains 1 DSL domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; I38493; AAB06569.1; -.

CC HSSP; P00740; 1EDM.

CC GO; GO:0005576; C:extracellular; ISS.

CC GO; GO:0005887; C:integral to plasma membrane; ISS.

CC GO; GO:0008083; F:growth factor activity; ISS.

CC GO; GO:0005112; F:Notch binding; ISS.

CC GO; GO:0005198; F:structural molecule activity; ISS.

CC GO; GO:0001525; F:angiogenesis; ISS.

CC GO; GO:0001709; P:cell fate determination; NAS.

CC GO; GO:0045446; P:endothelial cell differentiation; ISS.

CC GO; GO:0030097; P:hemoipoiesis; ISS.

CC GO; GO:0030216; P:keratinocyte differentiation; ISS.

CC GO; GO:004445; P:myoblast differentiation; ISS.

CC GO; GO:0007219; P:N signaling pathway; ISS.

CC GO; GO:0007399; P:neurogenesis; ISS.

CC DR: InterPro; IPR000152; P:regulation of cell proliferation; ISS.

CC DR: InterPro; IPR000152; Asx\_hydroxy1\_S.

CC DR: InterPro; IPR000174; DSL.

CC DR: InterPro; IPR000742; EGF.

CC DR: InterPro; IPR001881; EGF\_Ca.

CC DR: InterPro; IPR001438; EGF\_T1.

CC DR: InterPro; IPR005209; EGF\_Like.

CC DR: InterPro; IPR001007; WNF\_C.

CC DR: PFAM; PF0144; DSL\_1.

CC DR: PFAM; PF000008; EGF\_14.

CC DR: PRINTS; PR00010; EGF\_BLOOD.

CC DR: SMART; SMD0051; DSL\_1.

CC SMART; SMD0079; EGF\_Ca\_10.

CC SMART; SMD00214; VWC\_1.

CC DR: PROSITE; PS00010; ASX\_HYDROXYL\_10.

CC DR: PROSITE; PS00022; EGF\_1; 16.

CC DR: PROSITE; PS00186; EGF\_2; 12.

CC DR: PROSITE; PS00026; EGF\_3; 16.

CC DR: PROSITE; PS00187; EGF\_Ca\_8.

KW Calcium-binding; EGF-Like domain; Glycoprotein; Developmental protein; Repeat; Transmembrane; Signal.

PT SIGNAL\_1; SIGNAL\_33; POTENTIAL.

PT CHAIN\_34; JAGGED\_1.

PT DOMAIN\_34; EXTRACELLULAR (POTENTIAL).

PT TRANSMEM\_1068; POTENTIAL.

PT DOMAIN\_1094; CYTOPLASMIC (POTENTIAL).

PT DOMAIN\_167; DSL.

PT DOMAIN\_230; EGF-LIKE\_1 (ATYPICAL).

PT DOMAIN\_296; EGF-LIKE\_2.

PT DOMAIN\_336; EGF-LIKE\_3.

PT DOMAIN\_374; EGF-LIKE\_4; CALCIUM-BINDING (POTENTIAL).

PT DOMAIN\_410; EGF-LIKE\_5; CALCIUM-BINDING (POTENTIAL).

PT DOMAIN\_412; EGF-LIKE\_6; CALCIUM-BINDING (POTENTIAL).

PT DOMAIN\_450; EGF-LIKE\_7; CALCIUM-BINDING (POTENTIAL).

PT DOMAIN\_487; EGF-LIKE\_8.

PT DOMAIN\_525; EGF-LIKE\_9.

PT DOMAIN\_574; EGF-LIKE\_10; CALCIUM-BINDING (POTENTIAL).

PT DOMAIN\_629; EGF-LIKE\_11; CALCIUM-BINDING (POTENTIAL).

PT DOMAIN\_667; EGF-LIKE\_12.

PT DOMAIN\_705; EGF-LIKE\_13.

PT DOMAIN\_744; EGF-LIKE\_14; CALCIUM-BINDING (POTENTIAL).

PT DOMAIN\_782; EGF-LIKE\_15; CALCIUM-BINDING (POTENTIAL).

PT DISULFID\_300; BY SIMILARITY.

PT DISULFID\_306; BY SIMILARITY.

BY SIMILARITY.

324 FT DISULFID\_333; BY SIMILARITY.

340 FT DISULFID\_351; BY SIMILARITY.

345 FT DISULFID\_360; BY SIMILARITY.

362 FT DISULFID\_371; BY SIMILARITY.

378 FT DISULFID\_389; BY SIMILARITY.

383 FT DISULFID\_398; BY SIMILARITY.

400 FT DISULFID\_409; BY SIMILARITY.

416 FT DISULFID\_427; BY SIMILARITY.

421 FT DISULFID\_436; BY SIMILARITY.

438 FT DISULFID\_447; BY SIMILARITY.

454 FT DISULFID\_464; BY SIMILARITY.

458 FT DISULFID\_473; BY SIMILARITY.

475 FT DISULFID\_484; BY SIMILARITY.

491 FT DISULFID\_491; BY SIMILARITY.

496 FT DISULFID\_502; BY SIMILARITY.

511 FT DISULFID\_511; BY SIMILARITY.

513 FT DISULFID\_522; BY SIMILARITY.

529 FT DISULFID\_540; BY SIMILARITY.

534 FT DISULFID\_549; BY SIMILARITY.

551 FT DISULFID\_560; BY SIMILARITY.

578 FT DISULFID\_578; BY SIMILARITY.

599 FT DISULFID\_599; BY SIMILARITY.

617 FT DISULFID\_617; BY SIMILARITY.

633 FT DISULFID\_626; BY SIMILARITY.

638 FT DISULFID\_653; BY SIMILARITY.

655 FT DISULFID\_653; BY SIMILARITY.

671 FT DISULFID\_664; BY SIMILARITY.

671 FT DISULFID\_671; BY SIMILARITY.

676 FT DISULFID\_676; BY SIMILARITY.

676 FT DISULFID\_691; BY SIMILARITY.

693 FT DISULFID\_702; BY SIMILARITY.

709 FT DISULFID\_720; BY SIMILARITY.

714 FT DISULFID\_729; BY SIMILARITY.

731 FT DISULFID\_740; BY SIMILARITY.

748 FT DISULFID\_748; BY SIMILARITY.

753 FT DISULFID\_753; BY SIMILARITY.

770 FT DISULFID\_770; BY SIMILARITY.

786 FT DISULFID\_797; BY SIMILARITY.

791 FT DISULFID\_806; BY SIMILARITY.

808 FT DISULFID\_817; BY SIMILARITY.

824 FT DISULFID\_824; BY SIMILARITY.

829 FT DISULFID\_844; BY SIMILARITY.

846 FT DISULFID\_855; BY SIMILARITY.

855 FT DISULFID\_855; BY SIMILARITY.

855 FT CARBOYD\_143; N-LINKED (GLCNAC. . .) (POTENTIAL).

855 FT CARBOYD\_217; N-LINKED (GLCNAC. . .) (POTENTIAL).

855 FT CARBOYD\_382; N-LINKED (GLCNAC. . .) (POTENTIAL).

855 FT CARBOYD\_559; N-LINKED (GLCNAC. . .) (POTENTIAL).

855 FT CARBOYD\_745; N-LINKED (GLCNAC. . .) (POTENTIAL).

860 FT CARBOYD\_960; N-LINKED (GLCNAC. . .) (POTENTIAL).

891 FT CARBOYD\_991; N-LINKED (GLCNAC. . .) (POTENTIAL).

1045 FT CARBOYD\_1045; N-LINKED (GLCNAC. . .) (POTENTIAL).

1064 FT CARBOYD\_1064; N-LINKED (GLCNAC. . .) (POTENTIAL).

1219 AA; 134329 MW; SEQUENCE 1219 AA; 134329 MW; CRC84;

Query Match Score 86; DB 1; Length 1219;

Best Local Similarity 27.0%; Pred. No. 0.81; Gaps 9;

Matches 33; Conservative 9; Mismatches 34; Indels 46;

Qy 39 PE---GCTCEVIAAHRC---C-NKRNIBERSQTVKCSCLPGKVACTRNR-----81

Db 325 PEGYSNCE-TAEHAELSDPCHNRGSCKETSSGFBECSTINDDCSPPNC 383

Qy 82 ---PSCVDASIVG----KWW-----CEMPCLEGECKTLPDN-----SG 115

Db 384 SHGGTCQQ--LIVNGFKCVCPQWTGKTCOLDANECAKPCYNARSCKNLASYCDCLPG 441

Qy 116 WM 443

Db 442 WM 443

RESULT 13

ID FBPL\_STRPU STANDARD; PRT; 1064 AA.

AC P1.0079;

DT 01-MAR-1989 (Rel. 10, Created)



GN NOTCH4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates;  
NCBI TaxID=9606;  
[1] RN SEQUENCE FROM N.A. (ISOFORM 1), AND POLYMORPHISM OF POLY-LEU.  
RP TISSUE=Placenta;  
RC MEDLINE=97311416; PubMed=9168133;  
RX Sugaya K., Sasenuma S.-I., Nobata J., Kimura T., Fukagawa T., Nakamura Y., Ando A., Inoko H., Kimura K., Mita K.; RT "Gene organization of human NOTCH4 and (CTG)n Polymorphism in this human counterpart gene of mouse proto-oncogene Int3.";  
RT Gene 189:235-244 (1997).  
RN [2] SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RP TISSUE=Bone marrow, and Heart;  
RX MEDLINE=93360091; PubMed=1693032;  
RA Li L., Huang G.M., Banta A.B., Deng Y., Smith T., Dong P., Friedman C., Chen L., Trask B.J., Spies T., Rowen L., Hood L.; RT "Cloning, characterization, and the complete 56.8-kilobase DNA sequence of the human NOTCH4 gene.";  
RN Genomics 51:45-58(1998).  
[3] RN SEQUENCE OF 1-503 FROM N.A., AND VARIANTS GLN-117 AND GLN-317.  
RP RA Miyagawa T., Tokunaga K., Hojyo H.;  
RT "Human notch-4 gene variant."  
RL Submitted (FEB-1999) to the EMBL/GenBank/DDJBJ databases.  
RN [4] IDENTIFICATION OF LIGANDS.  
RP MEDLINE=99180765; PubMed=10079256;  
RX Gray G.E., Mann R.S., Mitsiadis E., Carcangiu M.-L., Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;  
RA RT "Human ligands of the Notch receptor.";  
RL Am. J. Pathol. 154:785-794 (1999).  
CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands Jagged1, Jagged2 and Delta1 to regulate cell-fate determination. Upon ligand activation through the released notch intracellular domain (NICD) it forms a transcriptional activator complex with RBP-J kappa and activates genes of the enhancer of split locus. Affects the implementation of differentiation, proliferation and apoptotic programs. May regulate branching morphogenesis in the developing vascular system (By similarity).  
CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-terminal fragment N(EC) which are probably linked by disulfide bonds (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following proteolytical processing NICD is translocated to the nucleus.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms-3;  
CC Comment=Experimental confirmation may be lacking for some isoforms;  
CC Name=1; ISOTId=Q99466-1; Sequence=Displayed;  
CC Name=2; ISOTId=Q99466-2; Sequence=YSP\_001406;  
CC Name=3; ISOTId=Q99466-3; Sequence=YSP\_001407;  
CC -!- CC TISSUE SPECIFICITY: Highly expressed in the heart, moderately in the lung and placenta and at low levels in the liver, skeletal muscle, kidney, pancreas, spleen, lymph node, thymus, bone marrow and fetal liver. No expression was seen in adult brain or peripheral blood leukocytes.  
CC -!- CC PTM: Synthesized in the endoplasmic reticulum as an inactive form which is proteolytically cleaved by a furin-like convertase in the trans-Golgi network before it reaches the plasma membrane to yield an active, ligand accessible form. Cleavage results in a C-terminal fragment N(TM) and a N-terminal fragment N(EC). Following ligand binding, it is cleaved by TNF-alpha converting enzyme (TACE) to yield a membrane-associated intermediate fragment called notch extracellular truncation (NEXT). This fragment is then cleaved by presenilin dependent gamma secretase to release a notch-derived peptide containing the intracellular domain (NICD).

	FT	DOMAIN	194	232	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	234	274	EGF-LIKE 6.
	FT	DOMAIN	276	312	EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	314	353	EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	335	391	EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	392	430	EGF-LIKE 10.
	FT	DOMAIN	432	473	EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	475	511	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	513	549	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	551	587	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	589	625	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
	FT	DOMAIN	626	659	EGF-LIKE 16.
	FT	DOMAIN	661	689	EGF-LIKE 17.
	FT	DOMAIN	691	727	EGF-LIKE 18.
	FT	DOMAIN	729	765	EGF-LIKE 19.
	FT	DOMAIN	767	803	EGF-LIKE 20.
	FT	DOMAIN	806	842	EGF-LIKE 21.
	FT	DOMAIN	844	880	EGF-LIKE 22.
	FT	DOMAIN	882	928	EGF-LIKE 23.
	FT	DOMAIN	930	966	EGF-LIKE 24.
	FT	DOMAIN	968	1004	EGF-LIKE 25.
	FT	DOMAIN	1006	1044	EGF-LIKE 26.
	FT	DOMAIN	1046	1085	EGF-LIKE 27.
	FT	DOMAIN	1087	1126	EGF-LIKE 28.
	FT	DOMAIN	1130	1171	POLY-LIKE 29.
	FT	REPEAT	1474	1476	POLY-ARG.
	FT	REPEAT	1165	1212	LIN/NOTCH 1.
	FT	REPEAT	1213	1246	LIN/NOTCH 2.
	FT	REPEAT	1247	1286	LIN/NOTCH 3.
	FT	REPEAT	1633	1665	ANK 1.
	FT	REPEAT	1666	1698	ANK 2.
	FT	REPEAT	1700	1732	ANK 3.
	FT	REPEAT	1733	1765	ANK 4.
	FT	REPEAT	1766	1798	ANK 5.
	FT	DISULFID	28	41	BY SIMILARITY.
	FT	DISULFID	35	51	BY SIMILARITY.
	FT	DISULFID	53	62	BY SIMILARITY.
	FT	DISULFID	68	80	BY SIMILARITY.
	FT	DISULFID	74	103	BY SIMILARITY.
	FT	DISULFID	105	114	BY SIMILARITY.
	FT	DISULFID	122	133	BY SIMILARITY.
	FT	DISULFID	127	143	BY SIMILARITY.
	FT	DISULFID	145	154	BY SIMILARITY.
	FT	DISULFID	160	171	BY SIMILARITY.
	FT	DISULFID	165	180	BY SIMILARITY.
	FT	DISULFID	182	191	BY SIMILARITY.
	FT	DISULFID	198	211	BY SIMILARITY.
Query Match			11.2%	Score 94.5; DB 1;	Length 2003;
Best Local Similarity			31.0%	Pred. No. 1.9;	
Matches			39;	Conservative 6; Mismatches	42; Indels 39; Gaps 10;
Qy	18	CAMLICH-GS	---	LOHTFQOQHHLRP	-EGGTCEV --- -IAAHRCCNKNLEERSQTVK 66
Dy	479	CLSQPCHGSSTCDLLAFT	---	HCLCPGCGLEGLCETVETNECASAPCLNAHDCHDINGFQ	536
Qy	67	CSCLPGKVAGTTNRPSC	---	YDASIVIGKWWCEMPEPLEGEE	----- CKTL,PDNS 114
Dy	537	CICLPG-FSGT	---	RCEEDIDE	----- CRSSPCANGQQCQDQGAFACKCJLPGEF 582
Qy	115	GMNCAT	120		
Dy	583	GPRCQT	588		

JAG2.  
 OS Mus musculus (Mouse).  
 OS Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus.  
 OC Submitted (DBC-1997) to the EMBL/GenBank/DDBJ databases.  
 OC [12]  
 RN SEQUENCE OF 302-819 FROM N.A.  
 RC TISSUE-Brain;  
 RX MEDLINE=98051908; PubMed=9341252;  
 RA Lan Y., Jiang R., Shawber C., Weinmaster G., Gridley T.;  
 RT "The Jagged2 gene maps to chromosome 12 and is a candidate for the lgl  
 and sm mutations";  
 RL Mamm. Genome 8:875-876 (1997).  
 RN SEQUENCE OF 325-759 FROM N.A.  
 RC TISSUE-Brain;  
 RX MEDLINE=98145947; PubMed=9486542;  
 RA Iuто B., Aster J.C., Hasserian R.P., Kuo F., Sklar J.;  
 RT "Isolation and functional analysis of a cDNA for human Jagged2, a gene  
 encoding a ligand for the Notch receptor.";  
 RT Mol. Cell. Biol. 17:6057-6067 (1997).  
 CC -!- FUNCTION: Putative Notch ligand involved in the mediation of Notch  
 signaling. Plays an essential role during limb, craniofacial and  
 thymic development. May be involved in myogenesis and in the  
 development of peripheral and central nervous systems.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Found to be highest in fetal thymus,  
 epidermis, foregut dorsal root ganglia and inner ear. In 2-week-old  
 mice, abundant in heart, lung, thymus, skeletal muscle, brain  
 and testis. Expression overlaps partially with Notch1 expression.  
 CC -!- DEVELOPMENTAL STAGE: At 13 dpc, found in paravertebral vessels and  
 dorsal root ganglia. At 14 dpc, in oropharyngeal epithelium,  
 developing thymus and in the muscles of the tongue. By 15 dpc, in  
 many tissues  
 CC -!- SIMILARITY: Contains 16 EGF-like domains.  
 CC -!- SIMILARITY: Contains 1 DSL domain.

---

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

---

CC EMBL; AF038572; AAF16411.1; -.  
 EMBL; AF010137; AAC14010.1; -.  
 DR HSSP; P00743; 1CCOF.  
 DR MGD; MGI:1098270; Jag2.  
 DR GO; GO:0005887; C:integral to plasma membrane; IDA.  
 DR GO; GO:0008033; F:growth factor activity; ISS.  
 DR GO; GO:0005112; F:Notch binding; IPI.  
 DR GO; GO:00107049; P:cell cycle; ISS.  
 DR GO; GO:0003054; P:cell differentiation; NAS.  
 DR GO; GO:0010109; P:cell fate determination; NAS.  
 DR GO; GO:0010767; P:cell-cell signaling; NAS.  
 DR GO; GO:0042492; P:gamma-delta T-cell differentiation; IMP.  
 DR GO; GO:0009912; P:hair cell fate commitment; IMP.  
 DR GO; GO:0010765; P:hair cell; hearing; NAS.  
 DR GO; GO:0030326; P:limb morphogenesis; ISS.

DR GO; GO:0007219; P:N signaling pathway; ISS.  
 DR GO; GO:0009334; P:regulation of cell migration; ISS.  
 DR GO; GO:0042127; P:regulation of cell proliferation; ISS.  
 DR GO; GO:0007283; P:spermatogenesis; ISS.  
 DR GO; GO:0045061; P:thymic T-cell selection; ISS.  
 DR InterPro; IPR000152; ASX hydroxy\_S.  
 DR InterPro; IPR001774; DSL.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR001881; EGF\_2A.  
 DR InterPro; IPR001438; EGF\_II.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR Pfam; PF01414; DSL; 1.  
 DR Pfam; PF00008; EGE; 14.  
 DR PRINTS; PRO0010; EGFBLOOD.  
 DR SMART; SM00051; DSL; 1.  
 DR SMART; SM00214; VWC; 1.  
 DR PROSITE; PS00022; ASX HYDROXYL; 10.  
 DR PROSITE; PS000179; EGF\_CA; 9.  
 DR SMART; SM00179; EGF\_CA; 9.  
 DR PROSITE; PS01186; EGF2; 16.  
 DR PROSITE; PS50028; EGF\_3; 15.  
 DR PROSITE; PS01187; EGF\_CA; 7.  
 KW Calcium-binding; EGF\_Like domain; Glycoprotein; Developmental protein;  
 KW repeat; Signal; Transmembrane.  
 SIGNAL 1 23  
 FT CHAIN 24 1247  
 FT DOMAIN 24 1084  
 FT TRANSMEM 24 1070  
 FT DOMAIN 1108 1247  
 FT DOMAIN 1178 240  
 FT DOMAIN 241 274  
 FT DOMAIN 275 305  
 FT DOMAIN 307 345  
 FT DOMAIN 347 383  
 FT DOMAIN 385 421  
 FT DOMAIN 423 459  
 FT DOMAIN 461 496  
 FT DOMAIN 498 534  
 FT DOMAIN 536 572  
 FT DOMAIN 574 634  
 FT DOMAIN 636 672  
 FT DOMAIN 674 710  
 FT DOMAIN 712 748  
 FT DOMAIN 751 787  
 FT DOMAIN 789 825  
 FT DOMAIN 827 863  
 FT DISULFID 245 256  
 FT DISULFID 249 262  
 FT DISULFID 264 273  
 FT DISULFID 276 287  
 FT DISULFID 282 293  
 FT DISULFID 295 304  
 FT DISULFID 311 323  
 FT DISULFID 317 333  
 FT DISULFID 335 344  
 FT DISULFID 351 362  
 FT DISULFID 356 371  
 FT DISULFID 373 382  
 FT DISULFID 389 400  
 FT DISULFID 394 409  
 FT DISULFID 411 420  
 FT DISULFID 427 438  
 FT DISULFID 432 447  
 FT DISULFID 449 458  
 FT DISULFID 465 475  
 FT DISULFID 469 484  
 FT DISULFID 486 495  
 FT DISULFID 502 513  
 FT DISULFID 507 522  
 FT DISULFID 524 533  
 FT DISULFID 540 551  
 FT DISULFID 560 560

FT DISULFID 562 BY SIMILARITY.  
 FT DISULFID 589 612 POTENTIAL.  
 FT DISULFID 606 622 BY SIMILARITY.  
 FT DISULFID 624 633 BY SIMILARITY.  
 FT DISULFID 640 651 BY SIMILARITY.  
 FT DISULFID 645 660 BY SIMILARITY.  
 FT DISULFID 662 671 BY SIMILARITY.  
 FT DISULFID 678 689 BY SIMILARITY.  
 FT DISULFID 683 698 BY SIMILARITY.  
 FT DISULFID 700 709 BY SIMILARITY.  
 FT DISULFID 716 727 BY SIMILARITY.  
 FT DISULFID 721 736 BY SIMILARITY.  
 FT DISULFID 738 747 BY SIMILARITY.  
 FT DISULFID 755 766 BY SIMILARITY.  
 FT DISULFID 760 775 BY SIMILARITY.  
 FT DISULFID 777 786 BY SIMILARITY.  
 FT DISULFID 793 804 BY SIMILARITY.  
 FT DISULFID 798 813 BY SIMILARITY.  
 FT DISULFID 815 824 BY SIMILARITY.  
 FT DISULFID 831 842 BY SIMILARITY.  
 FT DISULFID 836 851 BY SIMILARITY.  
 FT DISULFID 853 862 BY SIMILARITY.  
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 570 570 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 619 619 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1060 1060 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 302 302 L -> M (IN REF. 2).  
 FT CONFLICT 461 461 N -> T (IN REF. 2).  
 FT CONFLICT 469 478 CQRGGTCKDL -> VSANGHLQGP (IN REF. 2).  
 FT CONFLICT 492 492 G -> V (IN REF. 2).  
 FT CONFLICT 546 546 L -> F (IN REF. 2).  
 FT CONFLICT 549 549 A -> V (IN REF. 2).  
 FT CONFLICT 735 738 RCAC -> PAR (IN REF. 3).  
 FT CONFLICT 809 809 N -> H (IN REF. 2).  
 FT CONFLICT 812 812 R -> A (IN REF. 2).  
 SQ SEQUENCE 1247 AA: 134726 MW; 1D80C8626FAFABEC CRC64;  
 Qy CHGSLQHTQQHHLDHRPBGTC-EVIAAHC Score 84; DB 1; Length 1247;  
 Best Local Similarity 11.1%; Score 84; DB 1; Length 1247;  
 Matches 30; Conservative 12; Mismatches 30; Indels 50; Gaps 9;  
 Cnkrri- 58  
 Db 465 CHGSCQH-----GCTCDLNGYQCYPRGFGRHCELEYDKASSPBRGGIC 513  
 Qy 23 EERSQTVKCSCLPGKVAGTTRNREPSC-YDASIVIGKWCMEPCLEGECCKTLPDNSGWM 117  
 Qy 59 EERSQTVKCSCLPGKVAGTTRNREPSC-YDASIVIGKWCMEPCLEGECCKTLPDNSGWM 117  
 Db 514 EDLYDGFRCHC-PRGLSGL---HCEVDMDL----CEPSPLNGRCYND--EGDYY 559  
 Qy 118 CA 119  
 Db 560 CA 561

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 8, 2004, 13:15:04 ; Search time 22 Seconds  
(without alignments)  
312.102 Million cell updates/sec

Title: US-09-763-335-2  
Perfect score: 754  
Sequence: 1 MAMYSAMSIVYLWISACAM.....SGHMCATGNKIKTRIHPRT 133  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cggn2\_6/ptodata/2/iaa/5A\_COMB.pep:  
2: /cggn2\_6/ptodata/2/iaa/5B\_COMB.pep:  
3: /cggn2\_6/ptodata/2/iaa/6A\_COMB.pep:  
4: /cggn2\_6/ptodata/2/iaa/6B\_COMB.pep:  
5: /cggn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:  
6: /cggn2\_6/ptodata/2/iaa/backfile1.pep:  
\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	150	19.9	29	US-09-905-223-451	Sequence 451, App
2	110.5	14.7	1248	3	Sequence 6, App
3	96	12.7	1964	4	Sequence 1, App
4	91.5	12.1	233	4	Sequence 110, App
5	86.5	11.5	156	4	Sequence 22898, A
6	86.5	11.5	1139	1	Sequence 4, App
7	86.5	11.5	1139	3	Sequence 4, App
8	86.5	11.5	2703	1	Sequence 19, App
9	86.5	11.5	2703	4	Sequence 4, App
10	86	11.4	500	4	Sequence 2, App
11	86	11.4	659	4	Sequence 3, App
12	86	11.4	685	3	Sequence 2, App
13	86	11.4	685	3	Sequence 25, App
14	86	11.4	1010	3	Sequence 7, App
15	86	11.4	1036	4	Sequence 6, App
16	86	11.4	1187	4	Sequence 7, App
17	86	11.4	1208	4	Sequence 1, App
18	86	11.4	1218	2	Sequence 6, App
19	86	11.4	1218	3	Sequence 6, App
20	86	11.4	1218	3	Sequence 2, App
21	86	11.4	1218	3	Sequence 7, App
22	86	11.4	1218	4	Sequence 11, App
23	86	11.4	1218	4	Sequence 11, App
24	86	11.4	1219	3	Sequence 5, App
25	84	11.1	2523	1	Sequence 18, App
26	84	11.1	2523	4	Sequence 3, App
27	82.5	10.9	2471	1	Sequence 20, App

## ALIGNMENTS

● RESULT 1  
US-09-905-223-451  
; Sequence 451, Application US/08905223  
; Patent No. 6222029  
GENERAL INFORMATION:  
; APPLICANT: Edwards, Jean-Baptiste D.  
; APPLICANT: Dueiret, Aymeric  
; APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS: 503  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905, 223  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Israilev, Ned A.  
REGISTRATION NUMBER: 2, 655  
REFERENCE DOCUMENT NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 451:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE: Homo Sapiens  
ORGANISM: Homo Sapiens  
TISSUE/TYPE: Brain  
FEATURE:  
NAME/KEY: big peptide  
LOCATION: -25,-1  
IDENTIFICATION METHOD: Von Heijne matrix  
OTHER INFORMATION: score 5.8  
OTHER INFORMATION: seq LWISACMLCHG/SL  
US-08-905-223-451

Query Match 19.9%; Score 150; DB 3; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAMVSAMSWVLYLWISACMLLCHGSLQ 28  
 Db 1 MAMVSAMSWVLYLWISACMLLCHGSLQ 28

RESULT 2  
 US-08-882-046-6  
 / Sequence 6, Application US/08882046  
 / Patent No. 6136952  
 ; GENERAL INFORMATION:  
 / APPLICANT: Li, Linheng  
 / APPLICANT: Hood, Leroy  
 / APPLICANT: Krantz, Ian D.  
 / APPLICANT: Spinner, Nancy B.  
 / TITLE OF INVENTION: Human Jagged Polypeptide, Encoding  
 / TITLE OF INVENTION: Nucleic Acids and Methods of Use  
 / NUMBER OF SEQUENCES: 110  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: Campbell & Flores LLP  
 / STREET: 4370 La Jolla Village Drive, Suite 700  
 / CITY: San Diego  
 / STATE: California  
 / COUNTRY: USA  
 / ZIP: 92122  
 / COMPUTER READABLE FORM:  
 / COMPUTER: IBM PC compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: PatentIn Release #1.0, Version #1.25  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/08/882,046  
 / FILING DATE: 25-JUN-1997  
 / CLASSIFICATION: 536  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: Campbell, Kathryn A.  
 / REGISTRATION NUMBER: 31,815  
 / REFERENCE/DOCKET NUMBER: P-UW 2637  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: (619) 535-9001  
 / TELEFAX: (619) 535-8949  
 / INFORMATION FOR SEQ ID NO: 6:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 1248 amino acids  
 / TYPE:  
 / TOPOLOGY: Linear  
 / MOLECULE TYPE: peptide  
 US-08-882-046-6

Query Match 14.7%; Score 110.5; DB 3; Length 1248;  
 Best Local Similarity 27.0%; Pred. No. 0.0022;  
 Matches 33; Conservative 13; Mismatches 27; Indels 49; Gaps 9;

Qy 23 CHGSQLOHTEQQHHLRPEGSTC-EVIAAHR-C-----CNKNRI- 58  
 Db 465 CHGQCOH-----GETCKDLYNGYOCVCPRGFGGHCELEYKKCASSPCCRGGIC 513

Qy 59 EERSQTVKCSCLPGKVAGTRNRPSC-VDAASIVGKWWCMEPCLEGEEFKTLPDNSGMM 117  
 Db 514 EDLVDFRCH- PRGUSG----PLCEVDL-----WCPNPCTINGARCYLNED--YY 560

Qy 118 CA 119  
 Db 561 CA 562

; GENERAL INFORMATION:  
 / APPLICANT: Kitajewski, Jan  
 / APPLICANT: Uytterdael, Hendrik  
 / TITLE OF INVENTION: ANGIOTENSIN MODULATION BY NOTCH SIGNAL TRANSDUCTION  
 / FILE REFERENCE: 53863-A-PCT-US  
 / CURRENT APPLICATION NUMBER: US/09/467,997  
 / CURRENT FILING DATE: 1999-12-20  
 / NUMBER OF SEQ ID NOS: 10  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO: 1  
 ; LENGTH: 1964  
 ; TYPE: PRT  
 ; ORGANISM: mouse  
 US-09-467-997-1

Query Match 12.7%; Score 96; DB 4; Length 1964;  
 Best Local Similarity 26.8%; Pred. No. 0.13; Indels 34; Gaps 6;  
 Matches 33; Conservative 8; Mismatches 48;

Qy 15 ISACMLIC-HGSLOHTEQQHHLRPEGGTCA-----EVIAAHR--CCNKNRI 58  
 Db 669 LGCCISTPCAHGTCHP-----QPGYNTCTCPAGMGLTSEEVTAACHSGPCLNGGSC 741

Qy 59 EERSQTVKCSCLPGKVAGTRNRPSC-VDAASIVGKWWCMEPCLEGEEFKTLPDNSGMM 118  
 Db 742 SIRPEGYSTCTLP---SHTGRHQCTAVD-----HCVSASCLNGCTVNKPGTPCCLC 790

Qy 119 ATG 121  
 Db 791 ATG 793

RESULT 4  
 US-09-216-393B-110  
 / Sequence 110, Application US/09216393B  
 / Patent No. 6514694  
 ; GENERAL INFORMATION:  
 / APPLICANT: Milhauser, Michael James  
 / TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF  
 / FILE REFERENCE: TX-1-C2  
 / CURRENT APPLICATION NUMBER: US/09/216,393B  
 / PRIOR APPLICATION NUMBER: 08-12-18  
 / CURRENT FILING DATE: 1998-12-19  
 / PRIOR APPLICATION NUMBER: 08-994,825  
 / PRIOR FILING DATE: 1997-12-19  
 / NUMBER OF SEQ ID NOS: 366  
 / SOFTWARE: PatentIn version 3.1  
 / SEQ ID NO: 110  
 / LENGTH: 233  
 / TYPE: PRT  
 / ORGANISM: Toxoplasma gondii  
 US-09-216-393B-110

Query Match 12.1%; Score 91.5; DB 4; Length 233;  
 Best Local Similarity 25.9%; Pred. No. 0.036; Indels 35; Gaps 4;  
 Matches 28; Conservative 6; Mismatches 39;

Qy 18 CAMLICLGSQHTEQQHHLRPEGGTCA-----EVIAHRC-----CNKN-----RNBRPS 62  
 Db 78 CQCDACHG-----GKTCITKECCINCDSDCNGHGTNTNCNCEAGF 122

Qy 63 QTVKCSCLPGKVAGTRNRPSC-VDAASIVGKWWCMEPCLEGEEFKTL 110  
 Db 123 AGTNCSSSGKCSKTT-----CLSGHNPATGACVCDPCHTGERCETL 165

RESULT 5  
 US-09-252-991A-22898  
 / Sequence 22898, Application US/09252991A  
 / Patent No. 6551795  
 ; GENERAL INFORMATION:  
 / APPLICANT: Marc J. Rubenfield et al.  
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 / TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; Patent No. 6379925

RESULT 3  
 US-09-467-997-1  
 / Sequence 1, Application US/09467997  
 ; Patent No. 6379925

FILE REFERENCE: 107196-136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO: 22898  
 LENGTH: 156  
 TYPE: PRT  
 ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-22898

Query Match 11.5%; Score 86.5; DB 1; Length 1139;  
 Best Local Similarity 25.9%; Pred. No. 0.74; Gaps 3;  
 Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

Qy 40 EGCTGTYV---IAAHRCNKNRRIEERSQTVCSCLPKVA-----GTTNRPS 83  
 Db 26 QGQNCEINIDDCAPNPQNGTTCIDRMANFSCCPPTMGIICEINKDDCKPGACHING 85

Query Match 11.5%; Score 86.5; DB 4; Length 156;  
 Best Local Similarity 25.0%; Pred. No. 0.078; Gaps 5;  
 Matches 32; Conservative 9; Mismatches 64; Indels 23; Gaps 5;

Qy 6 AMSWVLWISACAMLLCHGSLOHTFOOHHLHRPEGGTCEVIAAHRCNKNRRIEERSQTVC 65  
 Db 12 ASW-XEWSAQ-TSTPGCLDTYAPARRPSARGSSKAVARRCCNATPCNSRSP 67

Qy 66 K-----CSCLPGKVAGTR----NRPSCVDASIVIGKVNCEMPECLEGEECKT 109  
 Db 68 NPCCLCWGSGSTCRLCLP1STDRMRPTASATRSTPSAVV--RWRBRPACTSMRCSR 124

Qy 110 LPDNGHM 117  
 Db 125 RSARRGWI 132

RESULT 6  
 US-08-537-210A-4  
 Sequence 4, Application US/08537210A  
 Patent No. 5780300  
 GENERAL INFORMATION:  
 APPLICANT: Aravaniis-Tsakonas, Spyridon  
 ATTORNEY/AGENT INFORMATION:  
 APPLICANT: Fortini, Mark  
 ATTORNEY/AGENT INFORMATION:  
 APPLICANT: Maruno, Kenji  
 TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY DIFFERENTIATED CELLS USING THE NOTCH PATHWAY  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10036/2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/537,210A  
 FILING DATE: 29-SEP-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Misrock, S. Leslie  
 REGISTRATION NUMBER: 18,872  
 REFERENCE/DOCKET NUMBER: 7326-027  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9050  
 TELEFAX: 212-869-8864  
 TELZEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1139 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein

Query Match 11.5%; Score 86.5; DB 3; Length 1139;  
 Best Local Similarity 25.9%; Pred. No. 0.74; Gaps 3;  
 Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

Qy 84 CYDASTIVIGKVNCEMPECLEGEECK 108  
 Db 86 CIDR--VGGFBCVCOQGFVGARCE 107

RESULT 7  
 US-09-113-825-4  
 Sequence 4, Application US/09113825  
 Patent No. 6149902  
 GENERAL INFORMATION:  
 APPLICANT: Aravaniis-Tsakonas, Spyridon  
 ATTORNEY/AGENT INFORMATION:  
 APPLICANT: Fortini, Mark  
 ATTORNEY/AGENT INFORMATION:  
 APPLICANT: Maruno, Kenji  
 TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY DIFFERENTIATED CELLS USING THE NOTCH PATHWAY  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10036/2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/113,825  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/537,210  
 FILING DATE: 29-SEP-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Misrock, S. Leslie  
 REGISTRATION NUMBER: 18,872  
 REFERENCE/DOCKET NUMBER: 7326-027  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9050  
 TELEFAX: 212-869-8864  
 TELZEX:  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1139 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein

NAME/KEY: Drss N  
 LOCATION: 1189--2327  
 OTHER INFORMATION: Highly conserved ankyrin repeat  
 US-08-537-210A-4

NAME/KEY: Drss N  
 LOCATION: 1189--2327  
 OTHER INFORMATION: Highly conserved ankyrin repeat  
 US-09-113-825-4

Query Match 11.5%; Score 86.5; DB 3; Length 1139;





REFERENCE/DOCKET NUMBER: P-UW 2637  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 535-9001  
 TELEFAX: (619) 535-8949  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1010 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-882-046-7

Query Match 11.4%; Score 86; DB 3; Length 1010;  
 Best Local Similarity 23.8%; Pred. No. 0.73;  
 Matches 34; Conservative 11; Mismatches 34; Indels 64; Gaps 9;  
 Qy 37 HRP-EGGTCE-----VIAAHRC---C-NKNRRIERSQTVKSC 69  
 Db 303 HQPCLNGTCNTGPDKYQCSCPEGYSSGPNCETAERHACLSDPCHNRGSKETSLGFRCBC 362  
 Qy 70 LPKVKAGTRNR-----PSCVDASIVG----KWW-----CEMEP 100  
 Db 363 SPWTGPTCSTNIDDCSPNNCSHGGTCQD-LVNGFKCVCPQWTGKNCQLDANECEAKP 420  
 Qy 101 CLEGECKTLPDN-----SGWM 117  
 Db 421 CYNAKSKCNLIAYYCDLPGM 443

RESULT 15  
 US-09-068-740A-6  
 Sequence 6, Application US/09068740A  
 Patent No. 6337387  
 GENERAL INFORMATION:  
 APPLICANT: SAKANO, SEIJI  
 APPLICANT: ITOH, AKIRA  
 TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE  
 CURRENT APPLICATION NUMBER: US/09/068,740A  
 CURRENT FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: JP 7-299611  
 PRIOR FILING DATE: 1995-11-17  
 PRIOR APPLICATION NUMBER: JP 7-311811  
 PRIOR FILING DATE: 1995-11-30  
 PRIOR APPLICATION NUMBER: PCT/JP96/03356  
 PRIOR FILING DATE: 1996-11-15  
 NUMBER OF SEQ ID NOS: 48  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 1036  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-068-740A-6

Query Match 11.4%; Score 86; DB 4; Length 1036;  
 Best Local Similarity 23.8%; Pred. No. 0.75;  
 Matches 34; Conservative 11; Mismatches 34; Indels 64; Gaps 9;  
 Qy 37 HRP-EGGTCE-----VIAAHRC---C-NKNRRIERSQTVKSC 69  
 Db 272 HQPCLNGTCNTGPDKYQCSCPEGYSSGPNCETAERHACLSDPCHNRGSKETSLGFRCBC 331  
 Qy 70 LPKVKAGTRNR-----PSCVDASIVG----KWW-----CEMEP 100  
 Db 332 SPWTGPTCSTNIDDCSPNNCSHGGTCQD-LVNGFKCVCPQWTGKNCQLDANECEAKP 389  
 Qy 101 CLEGECKTLPDN-----SGWM 117  
 Db 390 CYNAKSKCNLIAYYCDLPGM 412

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 8, 2004, 13:14:59 ; Search time: 20 Seconds

(without alignments)  
639 674 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MAMYSAMSATLYLNISACAM.....SGWMCATGNKIKTRIHPRT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters:

283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

PIR\_78:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	14.1	3871	2 T22812	hypothetical protein notch4 - mouse
2	96	12.7	1964	2 T09059	notch protein homo
3	95	12.7	2531	2 S18188	notch-1 protein -
4	95	12.7	2531	2 A46019	Notch homolog prot
5	94.5	12.5	2352	2 T30201	notch homolog - se
6	93	12.3	2531	2 T31072	crumbs protein - se
7	92.5	12.3	2139	2 A35672	notch protein - fr
8	86.5	11.5	2703	1 A24420	Delta-4 protein -
9	86	11.4	685	2 J570	jagged protein pre
10	86	11.4	1220	2 A56136	fibropallin-1a - s
11	85.5	11.3	1064	2 A40126	Xotch protein - Af
12	84	11.1	2524	2 S56058	Morph B protein -
13	82.5	10.9	1203	2 A49175	cell-fate determin
14	82.5	10.9	2471	2 A49128	notch 3 protein -
15	82	10.9	2321	2 S45306	alpha-2-macroglobu
16	82	10.9	293	2 B26637	zonadhesin - mouse
17	82	10.9	5376	2 T42215	probable membrane
18	81.5	10.8	355	2 S56058	notch protein homo
19	81.5	10.8	2555	2 A40043	hypothetical prote
20	81	10.7	907	2 T27317	notch3 protein - h
21	80.5	10.7	2321	2 S78549	neurogenic repeatit
22	79	10.5	4544	1 S02392	disintegrin and me
23	79	10.5	655	2 JC7850	metalloprotease
24	79	10.5	660	2 S71949	Noch homolog Motc
25	79	10.5	861	2 A48825	hypothetical prote
26	78.5	10.4	502	2 T20120	thrombospondin pre
27	78.5	10.4	838	2 T20125	MEGF2 protein - hu
28	78.5	10.4	1178	1 A39804	submitted to the EMBL Data Library, October 1997
29	78.5	10.4	1364	2 T00250	

## RESULT 1

T22812

hypothetical protein ZC116.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_change 15-Oct-1999

C;Accession: T22812; 127494

R;Burton, J.

submitted to the EMBL Data Library, June 1996

A;Reference number: Z19618

A;Accession: T22812

A;Status: preliminary; translated from GB/EMBL/DDJB

A;Molecule type: DNA

A;Residues: 1-3871 &lt;WIL&gt;

A;Cross-references: ENBL:274473; PIDN:CAA98952.1; GSPDB:GN00023; CESP:ZC116.3

A;Experimental source: Clone F56H9

R;Smee, R.

submitted to the EMBL Data Library, June 1996

A;Reference number: 220376

A;Accession: T27494

A;Cross references: ENBL:274046; PIDN:CAA98557.1; GSPDB:GN00023; CESP:ZC116.3

A;Experimental source: Clone ZC116

A;Genetics: C;Genetics:

A;Map position: 5

A;Introns: 29/3; 66/2; 244/1; 359/3; 422/1; 549/1; 572/1; 728/1; 820/1; 889/3; 1189/1; 1

C;Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology

Qy 24 HGSIQLTHQQHHHLHRPIEGTCVIAAPCCNNRERPSQ----TVYKCSCLPGKVAQT 77

Db 148 HQSL-----CQYIMSA-C-SKTFELCGPGHCICISIVDPTGQSSSDTTRYKCIIDBwgFcvss 202

Qy 78 TRURPSCYDASIVGKWCMEPCLEGBCKLPDN---SGWNMCATNK 123

Db 203 DRNPNTCVDN----ECESPCPHCPSVDCNCIPGSPVCSSG-CPKYK 243

Qy 24 HGSIQLTHQQHHHLHRPIEGTCVIAAPCCNNRERPSQ----TVYKCSCLPGKVAQT 77

Db 148 HQSL-----CQYIMSA-C-SKTFELCGPGHCICISIVDPTGQSSSDTTRYKCIIDBwgFcvss 202

Qy 78 TRURPSCYDASIVGKWCMEPCLEGBCKLPDN---SGWNMCATNK 123

Db 203 DRNPNTCVDN----ECESPCPHCPSVDCNCIPGSPVCSSG-CPKYK 243

C;Species: Mus musculus (house mouse)

C;Date: 11-Jun-1999 #sequence\_change 11-Jun-1999

C;Accession: T09059

R;Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.B.; Dankers, C.; Loretz, C.; Loretz, C.; Sc

submitted to the EMBL Data Library, October 1997

A;Description: Sequence of the mouse major histocompatibility locus class III region.

A;Reference number: Z16543

A;Accession: T0959

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1964 <ROW>

A;Cross-references: EMBL:AP030001; NID:g2564945; PMID:92564947

C;Genetics:

A;Gene: notchA

A;Map position: 17

A;Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 67

A;Exons: 1679/3; 1729/1; 1761/3

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

C;Keywords: receptor; signal transduction

F;514-545/Domain: EGF homology <EGF>

Query Match Score 96; DB 2; Length 1964;

Best Local Similarity 26.8%; Pred. No. 0.74;

Matches 33; Conservative 8; Mismatches 48; Indels 34; Gaps 6;

Db 689 ISACAMILCHGSLQHTPFOQHHLHRPEGTC -

Qy 59 ERSQTIVKSCCLPKVAGTTRNPRSPCVDAISIVGKRWCEMPLEGECKTLPDNSGPMC 118

Db 742 SIREPGYSCCLP -- SHTGBCQTAVD --- HCVSASCINGTGVNRPGTFFC 790

Qy 119 ATG 121

Db 791 ATG 793

## RESULT 3

S18188 notch protein homolog - rat

C;Species: Rattus norvegicus (Norway rat)

C;Accession: S18188 #sequence\_revision 10-Nov-1995 #text\_change 02-Aug-2002

R;Weinmaster, G.; Roberts, V.J.; Lemke, G.

A;Development: 11.1-205, 1991

A;Title: A homolog of Drosophila Notch expressed during mammalian development.

A;Reference number: S18188; MUID:92111383; PMID:1764995

A;Accession: S18188

A;Molecule type: mRNA

A;Residues: 1-2531 <WEL>

A;Cross-references: EMBL:X57405; NID:g57634; PMID:957635

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

F;987-1018/Domain: EGF homology <EGF1>

F;1025-1126/Domain: EGF homology <EGF2>

F;1233-1264/Domain: EGF homology <EGF3>

F;1917-1949/Domain: ankyrin repeat homology <AN1>

F;1950-1982/Domain: ankyrin repeat homology <AN2>

F;1984-2016/Domain: ankyrin repeat homology <AN3>

F;2017-2049/Domain: ankyrin repeat homology <AN4>

F;2050-2082/Domain: ankyrin repeat homology <AN5>

Query Match Score 95.5; DB 2; Length 2531;

Best Local Similarity 26.8%; Pred. No. 0.49;

Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

Db 679 IDECMGSPCH-----NGTCTEDGTAAGTCRCPGSYHDPTCLSEVNENCSN 723

Qy 57 -----RIEBSQTIVKSCCLPKVAGTTRNPRSPCVDAISIVGKRWCEMPLEBECKTLP 111

Db 724 PC1HGACRDGLNGYKCDARFG-WSGTNCD-----INNNCEESNPVNNGTKDM- 771

Qy 112 DNSGPMC 118

Db 772 -TSGIVC 777

## RESULT 4

A;6019

A;notch-1 protein - mouse

N;Alternative names: notch protein.

C;Species: Mus musculus (house mouse)

C;Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 07-Mar-2003

C;Accession: A46019; S25144; C9175; B4638; A46438; PM1569; S32109

R;del Amo, F.F.; Gentron-Maguire, M.; Jenkins, N.A.; Copeland, N.G.; Grid

Genomics 15, 259-264, 1993

A;Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

C;Keywords: receptor; signal transduction

F;514-545/Domain: EGF homology <EGF>

A;Description: Expression pattern of Notch, a mouse homolog of Drosophila Notch, suggest

A;Accession: A45019

A;Status: not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-2331 <DEL>

A;Cross-references: GB:547228; NID:g288502; PID:CAA77941.1; PMID:g288503

A;Note: sequence extracted from NCBI backbone (NCBIP:127318)

R;Franco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.;

submitted to the EMBL Data Library, April, 1992

A;Description: Expression pattern of Notch, a mouse homolog of Drosophila Notch, suggest

A;Accession: S25144

A;Molecule type: mRNA

A;Residues: 1551-2108, 'Q', 2110-2114, 'ALP', 2118-2170 <FRA>

A;Cross-references: EMBL:Z11886

A;Experimental source: embryo

A;Note: sequence extracted from NCBI backbone (NCBIP:126159)

R;Kopan, R.; Weintraub, H.; J. Cell Biol. 121, 651-641, 1993

A;Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety o

A;Reference number: A49175; MUID:93178563; PMID:93178563

A;Accession: C49175

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1161-1547 <LAR>

A;Cross-references: EMBL:X68278; NID:g287987; PID:CAA48339.1; PMID:g287988

A;Experimental source: embryo

A;Note: sequence extracted from NCBI backbone (NCBIP:126159)

J;Cell Biol. 104, 364-372, 1993

A;Title: Mouse notch expression in hair follicles correlates with cell fate determinati

A;Reference number: A46438; MUID:93252998; PMID:93252998

A;Accession: B46438

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1865-1932 /RR, 1935-1937 /I, 1938-1967 /I', 1969-2044 /IE, 2047-2052 /S, 2054

A;Experimental source: embryo

A;Note: sequence extracted from NCBI backbone (NCBIN:131246, NCBIP:131247)

C;Comment: This protein has many EGF repeats and lin-12(#1172) /Notch repeats.

C;Comment: This protein is one of the neurogenic proteins controlling the decision betwe

C;Genes: notch-1

A;Map position: 2

A;Note: proximal region of chromosome 2

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

F;106-139 /Domain: EGF homology <EGF1>

F;144-175 /Domain: EGF homology <EGF1>

F;222-254 /Domain: EGF homology <EGF2>

F;261-292 /Domain: EGF homology <EG02>

F;339-370 /Domain: EGF homology <EG03>

F;416-449 /Domain: EGF homology <EGF3>

F;456-487 /Domain: EGF homology <EG04>

F;532-563 /Domain: EGF homology <EG05>

F;607-638 /Domain: EGF homology <EG07>

F;782-813 /Domain: EGF homology <EG08>

F;757-788 /Domain: EGF homology <EG09>

F;795-826 /Domain: EGF homology <EG10>

F;813-840 /Domain: EGF homology <EG11>

F;941-942 /Domain: EGF homology <EG12>

F;949-980 /Domain: EGF homology <EG13>

F;987-1018 /Domain: EGF homology <EG14>

F;1025-1056 /Domain: EGF homology <EG15>

F;1063-1094 /Domain: EGF homology <EG16>

F;1149-1180/Domain: EGF homology <EG17>  
 F;1187-1218/Domain: EGF homology <EG18>  
 F;1264/Domain: EGF homology <EGF1>  
 F;1352-1383/Domain: EGF homology <EG19>  
 F;1391-1425/Domain: EGF homology <EG2>  
 F;1917-1948/Domain: ankyrin repeat homology <AN1>  
 F;1949-1981/Domain: ankyrin repeat homology <AN2>  
 F;1983-2015/Domain: ankyrin repeat homology <AN3>  
 F;2016-2049/Domain: ankyrin repeat homology <AN4>  
 F;2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 12.7%; Score 95.5; DB 2; Length 2531;  
 Best Local Similarity 26.8%; Pred. No. 0.49;  
 Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACAMILCHGSLQHTFQOHHLRPEGGTCE-VIAAHRC-----  
 DB 679 IDECAAGSPCH-----  
 QY 57 ---RIEERSQTVKXCSLPGKVAGTTRPSCVDASIVGKWCMEPCLEGEBCKTLR 111  
 DB 724 PCTHAGCRLGNGFKCDCAFP-WSGTNCD-----  
 QY 112 DNSGWMC 118  
 DB 772 -TSGYVC 777

RESULT 7  
 A35672 crumbs protein - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*  
 C;Date: 21-Sep-1990 #sequence\_revision 18-Nov-1992 #text\_change 21-Jul-2003  
 C;Accession: A35672  
 R;Repass', U.; Theres, C.; Knust, E.  
 Cell 61, 787-799, 1990  
 A;Title: crumbs encodes an EGF-like protein expressed on apical membranes of *Drosophila*  
 A;Reference number: A35672; PMID:90263104; PMID:2344615  
 A;Genetics:

A;Gene: FlyBase:crb  
 A;Cross-references: FlyBase:FBgn0003688  
 C;Keywords: transmembrane protein  
 A;Molecule type: mRNA  
 A;Residues: 1-2139 <TEP>  
 A;Cross-references: GB:M33753  
 A;Note: the authors translated the codon GCC for residue 1928 as Cys, and TAT for resid  
 A;Accession: A35672

T30201 Notch homolog protein - sea squirt (*Halocynthia roretzii*)

C;Accession: T30201 #sequence\_revision 02-Sep-2000 #text\_change 08-Sep-2002  
 R;Hori, S.; Saitoh, T.; Matsumoto, M.; Makabe, K.W.; Nishida, H.  
 Dev. Genes Evol. 207, 371-380, 1997  
 A;Title: Notch homologe from *Halocynthia roretzii* is preferentially expressed in the cen  
 A;Reference number: Z20075  
 A;Accession: T30201  
 A;Status: Preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-2352 <HOB>  
 A;Cross-references: EMBL:AB001327; NID:d1024472; PID:d1026501; PIDN:BAA25571.1  
 C;Genetics:  
 A;Gene: Notch protein; ankyrin repeat homology; EGF homology  
 C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match 12.5%; Score 94.5; DB 2; Length 2352;  
 Best Local Similarity 26.0%; Pred. No. 0.57;  
 Matches 27; Conservative 12; Mismatches 34; Indels 31; Gaps 5;

QY 32 QHHILHRPEGS-TCEVIAAH-----  
 DB 236 KRRD1QTEGFTCNVCYGYGTRDDSENIDCSNTACFANRCTDOAGTFECFLTPG-- 292

QY 76 GTTRNRPSCVDASIVGKWCMEPCLEGEBCKTLPDNSGWMC 118  
 DB 293 ---NRILCHDDA-----CISDPCARATCDINPITGSHWMC 325

RESULT 6  
 T31070 Notch homolog - sea urchin (*Lyttechinus variegatus*)  
 C;Species: Lyttechinus variegatus (variegated urchin)  
 C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 31-Jan-2000  
 C;Accession: T31070  
 R;Sherwood, D.R.; McClay, D.R.  
 Development 124, 3363-3374, 1997  
 A;Title: Identification and localisation of a sea urchin Notch homologue: insights into  
 A;Reference number: Z20066; NID:97454256; PMID:931031  
 A;Accession: T31070  
 A;Status: Preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA

QY 40 EGTCCEV---TAHRCCKNKRBERSTVKCSCLPGKVAGTTRNRSVC-----  
 DB 350 EGQTCTEINKDDTPNPQFGEGEDRVASFKCTCPGRTGLLCHLEACMSNPCHHPAQ 409

QY 86 DASTVIGKWKWCMEPCLEG-----  
 DB 410 STSVVDGSPICDATGGTQFGNCSEIDECISLSMDSTCOSGGTCQNFFGWSCLCSSGFIG 469

QY 122 NKTKT 126  
 DB 470 SRCT 474

RESULT 8  
 A24420 notch protein - fruit fly (*Drosophila melanogaster*)

N;Alternate names: neurogenic repetitive locus protein  
 C;Species: *Drosophila melanogaster*  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A24420; AA4768; S09458; A05267  
 R;Kidd, S.; Kelley, M.R.; Young, M.W.  
 Mol. Cell. Biol. 6, 3094-3108, 1986  
 A;Reference number: A24420  
 A;Accession: A24420  
 A;Molecule type: DNA  
 A;Residues: 1-2703 <KID>

A;Cross-references: GB:K03508; NID:g157991; PID:AAA28725.1; PID:g157993  
 R;Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.  
 Cell 143, 567-581, 1985  
 A;Reference number: A24768; MUID:86079539; PMID:3935325  
 A;Accession: A24768  
 A;Molecule type: mRNA  
 A;Residues: 1-48 'I', 50-118 'R', 120-230, 'I', 232-256 'N' 258-266 'A', 268-872 'R', 874-958;  
 A;Note: the authors translated the codon ATC for residue 49 as Thr, ATT for residue 2044  
 R;Whartuz, D.  
 Nucleic Acids Res. 17, 6463-6471, 1989  
 A;Title: Hypervariability of simple sequences as a general source for polymorphic DNA ma  
 A;Reference number: S0358; MUID:89385974; PMID:2780284  
 A;Accession: S0358  
 A;Molecule type: DNA  
 A;Residues: 2551-2551, 'QQQQ', 2552-2576, 'E', 2578-2604 <TAU,  
 R;Wharton, K.A.; Yedvobnick, B.; Finnerty, V.G.; Artavanis-Tsakonas, S.  
 Cell 40, 55-62, 1985  
 A;Title: opa: a novel family of transcribed repeats shared by the Notch locus and other  
 A;Reference number: A05267; MUID:85099329; PMID:2981631  
 A;Accession: A05267  
 A;Molecule type: DNA  
 A;Residues: 2504-2576, 'E', 2578-2611 <WHA2>  
 C;Genetics:  
 A;Gene: notch; QDA  
 A;Cross-references: FlyBase:FBgn004647  
 A;Map position: 8.96-9.36  
 A;Introns: 53/3; 94/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3  
 C;Superfamily: notch protein; ankyrin repeat homology; EGF homology  
 C;Keywords: differential tandem repeat; transmembrane protein  
 F;27-43:Domain: transmembrane #status predicted <TM1>  
 F;29-73:Domain: EGF homology <EGX1>  
 F;53-561/561:Domain: EGF homology <EGF1>  
 F;568-599/599:Domain: EGF homology <EGF2>  
 F;988-1019/1019:Domain: EGF homology <EGX2>  
 F;1064-1095/1095:Domain: EGF homology <EGF3>  
 F;1187-1218/1218:Domain: EGF homology <EGX3>  
 F;1746-1762/1762:Domain: transmembrane #status predicted <TM2>  
 F;1983-2015/2015:Domain: ankyrin repeat homology <AN1>  
 F;1988-2004/2004:Domain: transmembrane #status predicted <TM3>  
 F;2017-2049/2049:Domain: ankyrin repeat homology <AN3>  
 F;2050-2082/2082:Domain: ankyrin repeat homology <AN4>  
 F;2083-2115/2115:Domain: ankyrin repeat homology <AN5>  
 F;2538-2568/2568:Domain: neurogenic repetitive element #status predicted <OPA>  
 F;2569-2588/2588:Domain: neurogenic-rich

Query Match 11.5%; Score 86.5; DB 1; Length 2703;  
 Best Local Similarity 25.9%; Pred. No. 3.9;  
 Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

Qy 40 EGTCCEV---IAAHRCNNKRRIERERSQTVKCSCLPGKVA-----GTRTRRPS 83  
 Db 1214 QGONCELNIDCAPNPCCNGTCHDRVNNSCSCPPGTMGIICEINKDDCKPGACHNNGS 1273  
 Qy 84 CYDASITIVGKWNKEMEPGLEGECK 108  
 Db 1274 CIDR---VGGFFECVCQPRGVGARCE 1295

RESULT 9  
 JC75 70  
 Delta-4 protein - human  
 C;Species: Homo sapiens (mrat)  
 C;Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 16-Aug-2002  
 C;Accession: JC7570  
 R;Yoneyama, T.; Tahara, T.; Nagao, K.; Yamada, Y.; Yamamoto, T.; Osaka, M.; Miyatani, S.;  
 J. Biochem. 129, 27-34, 2001  
 A;Title: Molecular cloning of Delta-4, new mouse and human notch ligand.  
 A;Reference number: JC7569; MUID: 21064937; PMID:11134954  
 A;Accession: JC7570  
 A;Molecule type: mRNA  
 A;Residues: 1-685 <YON>  
 A;Cross-references: DDBJ:AB043894

Query Match 11.4%; Score 86; DB 2; Length 685;  
 Best Local Similarity 25.0%; Pred. No. 1.2;  
 Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

Qy 14 WISAC---AMLICHGSLOQHTFQQHHHLRPEGGTCEVIAHRCNNKRRIERERSQTVKCSC 69  
 Db 268 WQCTCDBGWGLFCDDQDNYC---THSPCRNGATC-----SNSGQSSEYT-CRC 312

Query Match 14 WISAC---AMLICHGSLOQHTFQQHHHLRPEGGTCEVIAHRCNNKRRIERERSQTVKCSC 69  
 Best Local Similarity 25.0%; Pred. No. 1.2;  
 Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

Qy 70 LPKGKVAGSTTRNRPSCTVADASITIVGKWNKEMEPGLEGECKTLPDNSGMCMATG 121  
 Db 313 RPYTG-----VOCLELSE--CDSNPGRNGGSCDKOEDGYHLCPPG 353

RESULT 10  
 A56136  
 Jagged protein precursor - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 21-Jul-2003  
 C;Accession: A56136  
 R;Lindell, C.E.; Shawber, C.J.; Boulter, J.; Weinmaster, G.  
 Cell 80, 909-917, 1995  
 A;Title: Jagged: a mammalian ligand that activates Notch1.  
 A;Reference number: A56136; MUID: 70597721  
 A;Accession: A56136  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-1220 <LIN>  
 A;Cross-references: GB:L38483  
 F;379-410/Domain: EGF homology <EGF1>  
 F;392-523/Domain: EGF homology <EGF2>  
 F;634-665/Domain: EGF homology <EGF2>  
 Query Match 11.4%; Score 86; DB 2; Length 1220;  
 Best Local Similarity 27.0%; Pred. No. 2.1;  
 Matches 33; Conservative 9; Mismatches 34; Indels 46; Gaps 9;

Qy 39 PE---GGTCCEVIAHRC---C-NKRNIEERSQTVKCSCLPGKVAGSTTRNR-----81  
 Db 326 PEGSGPNCE-TAEHALSLSDPHNRRGSSFEBCSPWTGPSCSTNLDDCSBNNC 384

Query Match 11.4%; Score 86; DB 2; Length 1220;  
 Best Local Similarity 27.0%; Pred. No. 2.1;  
 Matches 33; Conservative 9; Mismatches 34; Indels 46; Gaps 9;

Qy 82 ---PSCYDASITVG----KWW-----SG 115  
 Db 385 SHGGTCDD-LYNGFKCVCPPWTGKTCOLDANECPKCYNARSCKNLIASYYCDCLPG 442

Query Match 11.4%; Score 86; DB 2; Length 1220;  
 Best Local Similarity 27.0%; Pred. No. 2.1;  
 Matches 33; Conservative 9; Mismatches 34; Indels 46; Gaps 9;

Qy 116 WM 117  
 Db 443 WM 444

RESULT 11  
 A40136  
 Fibropellin Ia - sea urchin (Strongylocentrotus purpuratus)  
 N;Alternate names: epidermal growth factor homolog precursor  
 N;Contains: alternatively spliced fibropellin Ib (EGF1)  
 C;Species: Strongylocentrotus purpuratus (purple urchin)  
 C;Date: 13-May-1992 #sequence\_revision 17-Sep-1997 #text\_change 21-Jul-2000  
 C;Accession: A40136; B40136; C40136; A29316; A43131  
 R;Delgadillo-Reynoso, M.G.; Rollin, D.R.; Hurlsh, D.A.; Raff, R.A.  
 J. Mol. Evol. 29, 314-327, 1989  
 A;Title: Structural analysis of the UEGF gene in the sea urchin Strongylocentrotus purpu  
 A;Reference number: A40136; MUID: 9012459; PMID:2514273  
 A;Accession: A40136  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-114 <DEL>  
 A;Cross-references: GB:X17530; NID:910225; PID:g667061

R.Coffman, C.; Harris, W.; Kintner, C.	Science 249, 1438-1441, 1990	A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trs
A;Title: Xorch, the Xenopus homolog of Drosophila notch.		
A;Reference number: A35844; MUID:90388285; PMID:2402639		
A;Accession: A35844		
A;Molecule type: mRNA		
A;Residues: 1-2524 <COP>		
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology		
C;Keywords: transmembrane protein		
F:146-177/Domain: EGF homology <EGX1>		
F:184-215/Domain: EGF homology <BGF1>		
F:456-487/Domain: EGF homology <EGF2>		
F:757-788/Domain: EGF homology <EGX2>		
F:1025-1056/Domain: EGF homology <EGX3>		
F:1924-1956/Domain: ankyrin repeat homology <AN1>		
F:1957-1989/Domain: ankyrin repeat homology <AN2>		
F:1991-2023/Domain: ankyrin repeat homology <AN3>		
F:2024-2056/Domain: ankyrin repeat homology <AN4>		
F:2057-2089/Domain: ankyrin repeat homology <AN5>		
Query Match 11.1%; Score 84; DB 2; Length 2524;		
Best Local Similarity 23.4%; Pred. No. 6.4; Mismatches 41; Indels 8 Gaps 6;		
Matches 29; Conservative 16; Mi matches 1;		
Qy 23 CHGSLQHTPQQHHLHRPEGGTCEVLT----AAHRCCKNKRLEERS--QTUVKCSCLPGKVG 76		
Db 815 CNCMPPYT-----GAICBEAVLAPCGSPCKNGCRKCESEDPETFSCBPCPGMQGQ 864		
Qy 77 T-----TTRNPSCVDASIVIGRKCWCEMPICLEG-----EBCKTLPDNGSGWM 117		
Db 865 TCBDIMNCVYNRPCCRNGATCQNTN--GSYKCNCKFGYTGRCNCMDDDCOPENPNCHNGGS 921		
Qy 118 CATG 121		
Db 922 CSDG 925		
RESULT 13		
A49175	Motch B protein - mouse (fragment)	
N;Alternate names: Notch homolog		
C;Species: Mus musculus (house mouse)		
C;Date: 21-Jan-1996 #sequence revision 05-Jan-1996 #text_change 08-Sep-2002		
C;Accession: A49175; PMID: 832113		
R.Laruelle, M.; Lendahl, U.		
Exp. Cell Res. 204: 364-372, 1993		
A;Title: Moch C and Moch B -two mouse Notch homologues coexpressed in a wide variety of cells		
A;Reference number: A49175; MUID:3178563; PMID:840332		
A;Accession: A49175		
A;Status: preliminary; nucleic acid sequence not shown		
A;Molecule type: mRNA		
A;Residues: 1-1203 <LAR>		
A;Cross-references: EMBL_X68279; NID:9387989; PID:CAM48340_1; PID:9287990		
A;Experimental source: embryo		
A;Note: sequence extracted from NCBI backbone (NCBIP:126158)		
C;Comment: This protein has many EGF repeats and lin-12/Notch repeats.		
C;Comment: This protein is one of the neurogenin proteins controlling the decision between		
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology		
F:143-174/Domain: EGF homology <EGX1>		
F:432-511/Domain: EGF homology <EGF1>		
F:560-591/Domain: EGF homology <EGF2>		
F:674-705/Domain: EGF homology <EGX2>		
F:712-743/Domain: EGF homology <EGX3>		
F:836-867/Domain: EGF homology <EGX3>		
Query Match 10.9%; Score 82.5; DB 2; Length 1203;		
Best Local Similarity 26.4%; Pred. No. 4.5; Mismatches 11; Indels 17; Gaps 4;		
Matches 23; Conservative 11; Mi matches 36; Indels 17; Gaps 4;		
Qy 40 EGGTCEV----IAAHRCCKNKRLEERSQTVRCKSCLPGKVGTTTRNPSC-YDASIVGK 94		

Fri Apr 9 09:23:55 2004 /Domain: ankyrin repeat homology <AN5>

Query Match 10.9%; Score 82; DB 2; Length 2318;  
 Best Local Similarity 22.8%; Pred. No. 9.2.;  
 Matches 39; Conservative 16; Mismatches 42; Indels 74; Gaps 10;

QY 17 ACAM---LLCHGSLQHTFQOHHLHRP-----EGGTCE-----V 46  
 Db 340 ACPMGKTKLILCH-LDIDACVSNPCHEDAICCDTNPVSGRAICITCPGPGETGGACDQDVBCS 3.97

RESULT 14  
 A49128  
 cell-fate determining gene Notch2 protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 02-Aug-2002  
 C:Accession: A49128  
 R:Weimaster, G.: Roberts, V.J.; Lemke, G.  
 Development 116, 931-941, 1992  
 A>Title: Notch2: a second mammalian Notch gene.  
 A:Reference number: A49128; MUID:93202015; PMID:1295745  
 A:Accession: A49128  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-2471 <WETL>  
 A:Experimental source: Schwann cell  
 A:Note: sequence extracted from NCBI backbone (NCBIP:1.27811)  
 C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
 F:264-295/Domain: EGF homology <EGX1>  
 F:79-830/Domain: EGF homology <EGF1>  
 F:877-908/Domain: EGF homology <EGX2>  
 F:1029-1060/Domain: EGF homology <EGF2>  
 F:1067-1098/Domain: EGF homology <EGX3>  
 F:1153-1184/Domain: EGF homology <EGF3>  
 F:1191-1222/Domain: EGF homology <EGX4>  
 F:1876-1908/Domain: ankyrin repeat homology <AN1>  
 F:1909-1941/Domain: ankyrin repeat homology <AN2>  
 F:1943-1975/Domain: ankyrin repeat homology <AN3>  
 F:1976-2008/Domain: ankyrin repeat homology <AN4>  
 F:2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match 10.9%; Score 82.5%; DB 2; Length 2471;  
 Best Local Similarity 26.4%; Pred. No. 8.7%;  
 Matches 23; Conservative 11; Mismatches 36; Indels 17; Gaps 4;

QY 40 EGGTCEV----IAHRCCKNKRLEERSQTVKCSCLPGKVAGTTRNRPS-C-VDASIVGKW 94  
 Db 487 KGVHELEYNECQSNPCVNNQCVDKVNRFCQLCPGPFG-----PVQCIDDD---- 535

RESULT 15  
 S45306  
 notch 3 protein - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Aug-2002  
 C:Accession: S45306  
 R:Lardelli, M.; Dahlstrand, J.; Lendahl, U.  
 Mech. Dev. 46, 123-136, 1994  
 A>Title: The novel Notch homologue mouse Notch 3 lacks specific epidermal growth factor-  
 A:Reference number: S45306; MUID:95001556; PMID:7918097  
 A:Accession: S45306  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2318 <LAR>  
 A:Cross references: EMBL:X74760; MIM:9483580; PID:CAA52776.1; PID:9483581  
 C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
 F:163-195/Domain: EGF homology <EGF1>  
 F:474-505/Domain: EGF homology <EGF2>  
 F:554-585/Domain: EGF homology <EGF3>  
 F:1839-1871/Domain: ankyrin repeat homology <AN1>  
 F:1872-1904/Domain: ankyrin repeat homology <AN2>  
 F:1906-1938/Domain: ankyrin repeat homology <AN3>  
 F:1939-1971/Domain: ankyrin repeat homology <AN4>

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:17:35 ; Search time 40 Seconds  
(without alignments)  
874.269 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MARYSAMSWLYLMISACAM.....SGWMCATGNKIKTRIHPRT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1073127 seqs, 262937947 residues

Total number of hits satisfying chosen parameters: 1073127

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Published Applications AA:  
 1: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us07\_pubcomb.pep:  
 2: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us06\_pubcomb.pep:  
 3: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us05\_pubcomb.pep:  
 4: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us07\_new\_pub\_comb.pep:  
 5: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us06\_pubcomb.pep:  
 6: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us05\_pubcomb.pep:  
 7: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us08\_pubcomb.pep:  
 8: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us09\_pubcomb.pep:  
 9: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us09b\_pubcomb.pep:  
 10: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us08\_pubcomb.pep:  
 11: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us09c\_pubcomb.pep:  
 12: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us09\_new\_pub\_comb.pep:  
 13: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us10a\_pubcomb.pep:  
 14: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us10b\_pubcomb.pep:  
 15: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us10c\_pubcomb.pep:  
 16: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us10\_new\_pub\_comb.pep:  
 17: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us60\_new\_pub\_pep:  
 18: /cgm2\_6/\_ptodata/1/\_pubpaa/\_us60\_pubcomb.pep:  
 \* Pred. No. is the number of results predicted by score greater than or equal to the score of the result and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	106	14.1	3871	15	US-10-369-493-6538	Sequence 6538, AP
2	106	14.1	3871	15	US-10-369-493-6539	Sequence 6539, AP
3	95.5	12.7	2531	12	US-10-072-012-470	Sequence 470, APP
4	95.5	12.7	2531	12	US-10-072-012-471	Sequence 471, APP
5	95.5	12.7	2531	15	US-10-190-115-29	Sequence 29, APP
6	95.5	12.7	2531	15	US-10-369-072-29	Sequence 29, APP
7	91.5	12.1	233	9	US-09-216-39-110	Sequence 110, APP
8	91.5	12.1	233	14	US-10-321-856-110	Sequence 110, APP
9	88.5	11.7	2447	15	US-10-190-115-28	Sequence 28, APP
10	88.5	11.7	2447	15	US-10-369-072-28	Sequence 28, APP
11	87	11.5	188	9	US-09-764-853-773	Sequence 773, APP
12	87	11.5	188	9	US-09-764-898-262	Sequence 262, APP
13	87	11.5	188	10	US-09-764-881-109	Sequence 109, APP
14	87	11.5	188	14	US-10-073-865-88	Sequence 88, APP
15	87	11.5	188	15	US-10-242-747-109	Sequence 109, APP

RESULT 2

Organism: Caenorhabditis elegans

US-10-369-493-6538

Query Match Score 14.1%; Best Local Similarity 31.2%; Matches 34; Conservative 13; Mismatches 40; Indels 22; Gaps 6;

Prior Application Number: US 10/369 493 ; Current Application Number: US 10/369 493 ; Prior Filing Date: 2002-02-21 ; Number of SEQ ID NOS: 47374 ; SEQ ID NO: 6538 ; LENGTH: 3871 ; TYPE: PRT

FILE REFERENCE: 38-10152052.B ; CURRENT FILING DATE: 2003-02-28 ; PRIORITY FILING DATE: 2002-02-21 ; PRIORITY NUMBER: US 60/360,039

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES

GENERAL INFORMATION:

SEQUENCE: 6538, Application US/10369493 ; Publication No. US20030233675A1

APPLICATION: Cao, Yongwei ; Hinkle, Gregory J. ; Slater, Steven C. ; Goldman, Barry S. ; Applicant: Chen, Xianfang

INVENTOR: Chen, Xianfang

CURRENT APPLICATION NUMBER: US/10/369 493

US-10-369-493-6539  
*; Sequence 6539, Application US/10369493*  
*; Publication No. US20030233675A1*  
*; GENERAL INFORMATION:*  
*; APPLICANT: Cao, Yongwei*  
*; APPLICANT: Hinkle, Gregory J.*  
*; APPLICANT: Slater, Steven C.*  
*; APPLICANT: Goldman, Barry S.*  
*; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES*  
*; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES*  
*; FILE REFERENCE: 3B-10 (52052) B*  
*; CURRENT APPLICATION NUMBER: US/10/369,493*  
*; CURRENT FILING DATE: 2003-02-18*  
*; PRIOR APPLICATION NUMBER: US/10/369,039*  
*; NUMBER OF SEQ ID NOS: 47374*  
*; LENGTH: 3871*  
*; TYPE: PRT*  
*; ORGANISM: Caenorhabditis elegans*  
*US-10-369-493-6539*

Query Match 14.1%; Score 106; DB 15; Length 3871;  
 Best Local Similarity 31.2%; Pred. No. 0.15;  
 Matches 34; Conservative 13; Mismatches 40; Indels 22; Gaps 6;

Qy	24	HGSIQLQHTFOOHLRPEGCTCEVTAAHRCCKNKRIEERSQ-----TYSRSCSLPGRVAGT 77
Db	148	HGSI-----CQHNSAC-SKTFELPGPHGHCIESIVDPGQSSDDTTYKCICDWFGFYSS 202
Qy	78	TRNRPSCVDASTIVKGKWWCMEPCLEGECKTLPDN---SGNMCATGNK 123
Db	203	DKNRPTCVDN-----ECESNPCHPGVDCINLPGSFVCSG--CPKGYK 243

RESULT 3  
 US-10-072-012-470  
*; Sequence 470, Application US/10072012*  
*; Publication No. US20040033493A1*  
*; GENERAL INFORMATION:*  
*; APPLICANT: Tcherny, Velizar*  
*; APPLICANT: Spytek, Kimberly*  
*; APPLICANT: Zerhusen, Bryan*  
*; APPLICANT: Patturajan, Meera*  
*; APPLICANT: Shmikets, Richard*  
*; APPLICANT: Li, Li*  
*; APPLICANT: Gangolli, Esha*  
*; APPLICANT: Padigaru, Muralidhara*  
*; APPLICANT: Anderson, David W.*  
*; APPLICANT: Rastelli, Luca*  
*; APPLICANT: Shimkets, Richard*  
*; APPLICANT: Li, Li*  
*; APPLICANT: Gangolli, Esha*  
*; APPLICANT: Padigaru, Muralidhara*  
*; APPLICANT: Anderson, David W.*  
*; APPLICANT: Rastelli, Luca*  
*; APPLICANT: Miller, Charles E.*  
*; APPLICANT: Gerlach, Valerie*  
*; APPLICANT: Taupier Jr, Raymond J.*  
*; APPLICANT: Gusev, Vladimir Y.*  
*; APPLICANT: Colman, Steven D.*  
*; APPLICANT: Wolenc, Adam R.*  
*; APPLICANT: Pena, Carol E. A.*  
*; APPLICANT: Furtak, Katarzyna*  
*; APPLICANT: Grossbe, William M.*  
*; APPLICANT: Alsobrook II, John P.*  
*; APPLICANT: Lepley, Denise M.*  
*; APPLICANT: Rieger, Daniel K.*  
*; APPLICANT: Burgess, Catherine E.*  
*; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same*  
*; FILE REFERENCE: 21402-258*  
*; CURRENT APPLICATION NUMBER: US/10/072,012*  
*; CURRENT FILING DATE: 2002-01-31*  
*; PRIOR APPLICATION NUMBER: 60/265,102*  
*; PRIOR FILING DATE: 2001-01-30*  
*; PRIOR APPLICATION NUMBER: 60/255,514*  
*; PRIOR FILING DATE: 2001-01-31*  
*; PRIOR APPLICATION NUMBER: 60/265,517*

Query Match 12.7%; Score 95.5; DB 12; Length 2531;  
 Best Local Similarity 26.8%; Pred. No. 1.2;  
 Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

Qy	15	ISACAMILICHGSIQLQHTFOOHLRPEGCTCE-VIAAHRCC-----CNKN 56
Db	679	IDECASPH-----NGTCEDPGIAGETCRPEGYHADPTCLUSEVNBNNSN 723
Qy	57	-----REBERSQTVKCSLPLGKVAGTTRNRPSCVDASIVKGKWWCMEPCLEGECKTLP 111
Db	724	PCTHGAARDGLGYKCDCAPG-WSGTMC-----INNNECEENPCVNGGTCDM- 771
Qy	112	DNSGMMC 118
Db	772	-TSGYVC 777

RESULT 4  
 US-10-072-012-471  
*; Sequence 471, Application US/10072012*  
*; Publication No. US20040033493A1*  
*; GENERAL INFORMATION:*  
*; APPLICANT: Tcherny, Velizar*  
*; APPLICANT: Spytek, Kimberly*  
*; APPLICANT: Zerhusen, Bryan*  
*; APPLICANT: Patturajan, Meera*  
*; APPLICANT: Shmikets, Richard*  
*; APPLICANT: Li, Li*  
*; APPLICANT: Gangolli, Esha*  
*; APPLICANT: Padigaru, Muralidhara*  
*; APPLICANT: Anderson, David W.*  
*; APPLICANT: Rastelli, Luca*  
*; APPLICANT: Miller, Charles E.*  
*; APPLICANT: Gerlach, Valerie*  
*; APPLICANT: Taupier Jr, Raymond J.*  
*; APPLICANT: Gusev, Vladimir Y.*  
*; APPLICANT: Colman, Steven D.*  
*; APPLICANT: Wolenc, Adam R.*  
*; APPLICANT: Pena, Carol E. A.*  
*; APPLICANT: Furtak, Katarzyna*  
*; APPLICANT: Grossbe, William M.*  
*; APPLICANT: Alsobrook II, John P.*  
*; APPLICANT: Lepley, Denise M.*  
*; APPLICANT: Rieger, Daniel K.*  
*; APPLICANT: Burgess, Catherine E.*  
*; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same*  
*; FILE REFERENCE: 21402-258*  
*; CURRENT APPLICATION NUMBER: US/10/072,012*  
*; CURRENT FILING DATE: 2002-01-31*  
*; PRIOR APPLICATION NUMBER: US/10/072,012*  
*; CURRENT FILING DATE: 2002-01-31*

PRIOR APPLICATION NUMBER: 60/265,102  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: 60/265,514  
 PRIOR FILING DATE: 2001-01-31  
 PRIOR APPLICATION NUMBER: 60/265,517  
 PRIOR FILING DATE: 2001-01-31  
 PRIOR APPLICATION NUMBER: 60/265,412  
 PRIOR FILING DATE: 2001-01-31  
 PRIOR APPLICATION NUMBER: 60/265,395  
 PRIOR FILING DATE: 2001-01-31  
 PRIOR APPLICATION NUMBER: 60/266,406  
 PRIOR FILING DATE: 2001-02-02  
 PRIOR APPLICATION NUMBER: 60/266,767  
 PRIOR FILING DATE: 2001-02-05  
 PRIOR APPLICATION NUMBER: 60/267,057  
 PRIOR FILING DATE: 2001-02-07  
 PRIOR APPLICATION NUMBER: 60/266,975  
 PRIOR FILING DATE: 2001-02-07  
 PRIOR APPLICATION NUMBER: 60/267,459  
 PRIOR FILING DATE: 2001-02-08  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 1391  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 471  
 LENGTH: 2531  
 TYPE: PRT  
 ORGANISM: Rattus norvegicus  
 US-10-072-012-471

Query Match 12.7%; Score 95.5; DB 12; Length 2531;  
 Best Local Similarity 26.8%; Pred. No. 1.2; Mismatches 7;  
 Matches 34; Conservative 7; Indels 51; Gaps 7;  
 US-10-190-115-29

Qy 15 ISACAMILCHGSLOHTPQOHHLARPEGGTE-VIAAHC-  
 Db 679 IDCAGSPCH-----NGGTCTEDGTAAGTICRCPEGYHDPTCLSEVNNECSNSN 723  
 Qy 57 ----RIBERSQTVKCSLCPKGVAGTTRNRPSCVDASIVIGKWNCEMPCLEGEBCCKTLP 111  
 Db 724 PCTHGA CRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVCVNGGTCKDM- 771

Qy 112 DNSGMWC 118  
 Db 772 -TSGYVC 777

RESULT 5  
 US-10-190-115-29  
 Sequence 29, Application US/10190115  
 Publication No. US20030207394A1  
 GENERAL INFORMATION:

APPLICANT: Alsobrook, John P. II  
 APPLICANT: Boldog, Ferenc L.  
 APPLICANT: Burgess, Catherine E.  
 APPLICANT: Casman, Stacie J.  
 APPLICANT: Gross, William M.  
 APPLICANT: Gusev, Vladimir Y.  
 APPLICANT: Ji, Weizhen  
 APPLICANT: Lepley, Denise M.  
 APPLICANT: Liu, Xiaohong  
 APPLICANT: Nezick, Amanda J.  
 APPLICANT: Padigaru, Muralidhara  
 APPLICANT: Paturajan, Meera  
 APPLICANT: Rastelli, Luca  
 APPLICANT: Shen, Lei  
 APPLICANT: Shenoy, Suresh G.  
 APPLICANT: Shimkens, Richard A.  
 APPLICANT: Spaderna, Steven K.  
 APPLICANT: Spytek, Kimberly A.  
 APPLICANT: Szekeres, Edward S. Jr.  
 APPLICANT: Taupier, Raymond J. Jr.  
 APPLICANT: Tchernev, Velizar T.  
 APPLICANT: Zerhusen, Bryan D.

APPLICANT: Voss, Edward Z.  
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 FILE REFERENCE: 21402-050 CIP  
 CURRENT APPLICATION NUMBER: US/10/190,115  
 CURRENT FILING DATE: 2003-02-10  
 PRIOR APPLICATION NUMBER: 60/303,168  
 PRIOR FILING DATE: 2001-07-05  
 PRIOR APPLICATION NUMBER: 60/368,996  
 PRIOR FILING DATE: 2002-04-01  
 PRIOR APPLICATION NUMBER: 60/386,816  
 PRIOR FILING DATE: 2002-06-07  
 PRIOR APPLICATION NUMBER: 60/215,854  
 PRIOR FILING DATE: 2000-07-03  
 PRIOR APPLICATION NUMBER: 60/215,856  
 PRIOR FILING DATE: 2000-07-03  
 PRIOR APPLICATION NUMBER: 60/215,902  
 PRIOR FILING DATE: 2000-07-03  
 PRIOR APPLICATION NUMBER: 60/216,585,  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: 60/216,586  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: 60/216,722  
 PRIOR FILING DATE: 2000-07-07  
 PRIOR APPLICATION NUMBER: 60/218,622  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 136  
 SOFTWARE: CuraseqList version 0.1  
 SEQ ID NO: 29  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-190-115-29

Query Match 12.7%; Score 95.5; DB 15; Length 2531;  
 Best Local Similarity 26.8%; Pred. No. 1.2;  
 Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;  
 US-10-190-115-29

Qy 15 ISACAMILCHGSLOHTPQOHHLARPEGGTE-VIAAHC-  
 Db 679 IDECAGSPCH-----NGGTCTEDGTAAGTICRCPEGYHDPTCLSEVNNECSNSN 723  
 Qy 57 ----RIBERSQTVKCSLCPKGVAGTTRNRPSCVDASIVIGKWNCEMPCLEGEBCCKTLP 111  
 Db 724 PCTHGA CRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVCVNGGTCKDM- 771

Qy 112 DNSGMWC 118  
 Db 772 -TSGYVC 777

RESULT 6  
 US-10-369-072-29  
 Sequence 29, Application US/10369072  
 Publication No. US2004001408A1  
 GENERAL INFORMATION:  
 APPLICANT: Alsobrook II, John P.  
 APPLICANT: Spaderna, Stephen K.  
 APPLICANT: Tchernev, Velizar  
 APPLICANT: Liu, Xiaohong  
 APPLICANT: Shenoy, Suresh  
 APPLICANT: Spytek, Kimberly  
 APPLICANT: Zerhusen, Bryan  
 APPLICANT: Paturajan, Meera  
 APPLICANT: Rastelli, Luca  
 APPLICANT: Taupier, Raymond T.  
 APPLICANT: Bastelli, Luca  
 APPLICANT: Gross, William M.  
 APPLICANT: Szekeres, Edward S.  
 APPLICANT: Shen, Lei  
 APPLICANT: Burgess, Catherine E.  
 APPLICANT: Shimkens, Richard A.  
 APPLICANT: Spaderna, Steven K.  
 APPLICANT: Spytek, Kimberly A.  
 APPLICANT: Szekeres, Edward S. Jr.  
 APPLICANT: Taupier, Raymond J. Jr.  
 APPLICANT: Tchernev, Velizar T.  
 APPLICANT: Zerhusen, Bryan D.

TITLE OF INVENTION: No. US20040014081A1; Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21102-050 CON2  
; CURRENT FILING DATE: US/10/369,072  
; PRIOR APPLICATION NUMBER: 10/03-02-18  
; PRIOR APPLICATION NUMBER: 10/174,372  
; PRIOR APPLICATION NUMBER: 10/02-16-17  
; PRIOR APPLICATION NUMBER: 09/898,994  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: 60/215,054  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: 60/215,856  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: 60/215,902  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: 60/216,585  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/216,586  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/216,722  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/218,622  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 60/218,992  
; PRIOR FILING DATE: 2000-07-17  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 29  
; LENGTH: 2531  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; US-10-369-072-29

Query Match Score 95.5%; DB 15; Length 2531;  
Best Local Similarity 26.8%; Pred. No. 1.2;  
Matches 34; Conservative 7; Mismatches 35;  
Indels 51; Gaps 7;

Qy 15 ISACMLLGSLSLQHTFQOHLRHPGGTCE-VIAHRC-----CNKN 56  
Db 679 IDEAGSPCH-----NGGCTDGIAGFTCRCPGYHDPTCLSEVNCCNSN 723  
Qy 57 -----RIER-SOTYRKCSCLPGKVAGTTRNRPSCWDASIVIGKWCMEPCLGEECKTL 111  
Db 724 PCIHACRQDGNLGYRCDARG-WSGTNCD-----INNECESNVCNGTCKDM- 771  
Qy 112 DNSGMWC 11.8  
Db 772 -TSGYVC 777

Query Match Score 91.5%; DB 14; Length 233;  
Best Local Similarity 25.9%; Pred. No. 0.23;  
Matches 28; Conservative 6; Mismatches 39;  
Indels 35; Gaps 4;

Qy 18 CAMLLCGSLOHTFQOHLRHPGGTCEVIAHRC-----CNKN 62  
Db 78 CQCDACHG-----GKTEITKEHCCINDSDCNGHGTNTSNNTCNEAGF 122  
Qy 63 QTVKCSCLPGKVAGTTRNRPSCWDASIVIGKWCMEPCLGEECKTL 110  
Db 123 AGTNCSSSEKGCSGKT-----CLSGHNPATGACVCPCHTCERCEL 165

RESULT 7  
US-09-216-393-110  
; Sequence 11.0, Application US/09216393  
; Patent No. US2004001447A1  
; GENERAL INFORMATION:  
; APPLICANT: Milhausen, Michael James  
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: TX-1-C  
; CURRENT FILING DATE: US/09/216,393  
; EARLIER APPLICATION NUMBER: 1998-12-18  
; EARLIER FILING DATE: 1997-12-19  
; NUMBER OF SEQ ID NOS: 364  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 11.0  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Toxoplasma gondii  
; US-09-216-393-110

Query Match Score 91.5%; DB 9; Length 233;  
Best Local Similarity 25.9%; Pred. No. 0.23;

RESULT 8  
US-10-321-856-110  
; Sequence 11.0, Application US/10321856  
; Publication No. US2003019439A1  
; GENERAL INFORMATION:  
; APPLICANT: Milhausen, Michael James  
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF  
; FILE REFERENCE: TX-1-C1-1  
; CURRENT APPLICATION NUMBER: US/10/321-856  
; CURRENT FILING DATE: 2002-12-17  
; PRIOR APPLICATION NUMBER: 09/216,393  
; PRIORITY FILING DATE: 1998-12-18  
; PRIOR APPLICATION NUMBER: 08/994,825  
; PRIORITY FILING DATE: 1997-12-19  
; NUMBER OF SEQ ID NOS: 366  
; SOFTWARE: PatentIn Version 3.1  
; SEQ ID NO: 110  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Toxoplasma gondii  
; US-10-321-856-110

Query Match Score 91.5%; DB 14; Length 233;  
Best Local Similarity 25.9%; Pred. No. 0.23;  
Matches 28; Conservative 6; Mismatches 39;  
Indels 35; Gaps 4;

Qy 18 CAMLLCGSLOHTFQOHLRHPGGTCEVIAHRC-----CNKN 62  
Db 78 CQCDACHG-----GKTEITKEHCCINDSDCNGHGTNTSNNTCNEAGF 122  
Qy 63 QTVKCSCLPGKVAGTTRNRPSCWDASIVIGKWCMEPCLGEECKTL 110  
Db 123 AGTNCSSSEKGCSGKT-----CLSGHNPATGACVCPCHTCERCEL 165

RESULT 9  
US-10-190-115-28  
; Sequence 28, Application US/10190115  
; Publication No. US20030207394A1  
; GENERAL INFORMATION:  
; APPLICANT: Alabrook, John P. II  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Burgess, Catherine E.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Gross, William M.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Leyley, Denise M.  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Mezick, Amanda J.  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Paturajan, Meera  
; APPLICANT: Bastelli, Luca  
; APPLICANT: Shen, Lei  
; APPLICANT: Shanoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Stekeris, Edward S. Jr.  
; APPLICANT: Taupier, Raymond J. Jr.  
; APPLICANT: Tcherniev, Velizar T.  
; APPLICANT: Zerhusen, Bryan D.

APPLICANT: Voss, Edward Z.  
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 FILE REFERENCE: 2142-050 CIP  
 CURRENT APPLICATION NUMBER: US/10/190,115  
 CURRENT FILING DATE: 2003-02-10  
 PRIORITY NUMBER: 60/303,168  
 PRIOR FILING DATE: 2001-07-05  
 PRIORITY NUMBER: 60/368,996  
 PRIOR FILING DATE: 2000-04-01  
 PRIORITY NUMBER: 60/386,816  
 PRIOR FILING DATE: 2002-06-07  
 PRIORITY NUMBER: 60/215,854  
 PRIOR FILING DATE: 2000-07-03  
 PRIORITY NUMBER: 60/215,856  
 PRIOR FILING DATE: 2000-07-03  
 PRIORITY NUMBER: 60/215,902  
 PRIOR FILING DATE: 2000-07-03  
 PRIORITY NUMBER: 60/216,585  
 PRIOR FILING DATE: 2000-07-07  
 PRIORITY NUMBER: 60/216,586  
 PRIOR FILING DATE: 2000-07-03  
 PRIORITY NUMBER: 60/216,586  
 PRIOR FILING DATE: 2000-07-03  
 PRIORITY NUMBER: 60/215,902  
 PRIOR FILING DATE: 2000-07-03  
 PRIORITY NUMBER: 60/218,622  
 PRIOR FILING DATE: 2000-07-17  
 PRIORITY NUMBER: 60/218,992  
 PRIOR FILING DATE: 2000-07-17  
 PRIORITY NUMBER: 60/218,992  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 SEQ ID NO: 28  
 SOFTWARE: OuraSeqList version 0.1  
 LENGTH: 2447  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 SEQ ID NO: 136  
 NUMBER OF SEQ ID NOS: 136  
 SEQ ID NO: 28  
 LENGTH: 2447  
 PRIORITY NUMBER: 60/216,722  
 PRIOR FILING DATE: 2000-07-07  
 PRIORITY NUMBER: 60/218,622  
 PRIOR FILING DATE: 2000-07-17  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 SEQ ID NO: 28  
 LENGTH: 2447  
 PRIORITY NUMBER: 60/216,722  
 PRIOR FILING DATE: 2000-07-07  
 PRIORITY NUMBER: 60/218,622  
 PRIOR FILING DATE: 2000-07-17  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 SEQ ID NO: 28  
 LENGTH: 2447  
 PRIORITY NUMBER: 60/216,722  
 PRIOR FILING DATE: 2000-07-07  
 PRIORITY NUMBER: 60/218,622  
 PRIOR FILING DATE: 2000-07-17  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 SEQ ID NO: 28  
 LENGTH: 2447  
 PRIORITY NUMBER: 60/216,722  
 PRIOR FILING DATE: 2000-07-07  
 PRIORITY NUMBER: 60/218,622  
 PRIOR FILING DATE: 2000-07-17  
 Remaining Prior Application data removed - See File Wrapper or PALM.

Query Match 11.7%; Score 88.5; DB 15; Length 2447;  
 Best Local Similarity 25.4%; Pred. No. 5.9; Mismatches 39; Indels 39; Gaps 5;

Qy 41 GGTCT---VIAAHRCNCNRIBERSQTVKCSCLPGKVAGT-----TRNPF 83  
 Db 404 GPRCEQDYNEMECAASNPQNQDTCLDRIGDYSICMPG-FGGTH-BENELNBCLSSPCLNRGK 462

Qy 84 CVDASIVIGKWWCE-----MEPCLEGECKTLPDNSGWMCATGNK 123  
 Db 463 CLDQ---VSRFVCCECPAGEGEMCQIDDECSCSPCLNRGK 517

RESULT 11  
 US-09-764-853-773  
 Sequence 773, Application US/09764853  
 Patent No. US20020090672A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PTO6  
 CURRENT APPLICATION NUMBER: US/09/764,853  
 CURRENT FILING DATE: 2001-01-17  
 Prior application data removed - consult PALM or file wrapper  
 NUMBER OF SEQ ID NOS: 939  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 773  
 LENGTH: 188  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (97)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (187)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (188)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-764-853-773

Query Match 11.5%; Score 87; DB 9; Length 188;  
 Best Local Similarity 25.0%; Pred. No. 0.53;

GENERAL INFORMATION:  
 APPLICANT: Alsobrook II, John P  
 APPLICANT: Spadera, Stephen K  
 APPLICANT: Tchernev, Velizar  
 APPLICANT: Liu, Xiaohong  
 APPLICANT: Shanyo, Suresh  
 APPLICANT: Spytek, Kimberly  
 APPLICANT: Zerhusen, Bryan  
 APPLICANT: Patturajan, Meera  
 APPLICANT: Taupier, Raymond T  
 APPLICANT: Rastelli, Luca  
 APPLICANT: Gross, William M  
 APPLICANT: Sperkeres, Edward S  
 APPLICANT: Lepley, Denise M  
 APPLICANT: Shen, Lei  
 APPLICANT: Burgess, Catherine E  
 APPLICANT: Shimkets, Richard  
 APPLICANT: Padigaru, Muralidhara  
 APPLICANT: Title of Invention: No. US200401081A1  
 Proteins and Nucleic Acids Encoding Same  
 FILE REFERENCE: 2142-050 CON2  
 CURRENT APPLICATION NUMBER: US/10/369,072  
 CURRENT FILING DATE: 2003-02-18

Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Qy 14 WISAC----AMLLCHGSLOFTFOOFLHREGGTCEVIAHRCNCNKRTEERSQTVKCSC 69 NAME/KEY: SITE  
; LOCATION: (187)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Db 18 WQCTCDEGGGLFCDDLNIC--THSPCKNGATC-----SNSGQRSYT--CTC 62 NAME/KEY: SITE  
; LOCATION: (188)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Qy 70 LPGKVAGTTRNRPSCYDASIVGKWCMEPCLGECKTLPDNSGWMCATG 121 US-09-764-881-109  
; LENGTH: 122

Db 63 RPGYTG-----VDCEELSE--CDSNPCRNGGSKCQEDGYXCLCPG 103

Query Match 11.5%; Score 87; DB 10; Length 188;  
Best Local Similarity 25.0%; Pred. No. 0.53;  
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

Qy 14 WISAC----AMLLCHGSLOHTFOOHLHRPEGGTCEVIAAHRCNCNKRLEERSQTVKCSC 69  
; LENGTH: 122  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Db 18 WQCTCDEGGGLFCDDLNIC--THHSPOCKNGATC-----SNSGQRSYT--CTC 62  
; LENGTH: 122

Qy 70 LPGKVAGTTRNRPSCYDASIVGKWCMEPCLGECKTLPDNSGWMCATG 121  
; LENGTH: 122  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Db 63 RPGYTG-----VDCEELSE--CDSNPCRNGGSKCQEDGYXCLCPG 103  
; LENGTH: 122

RESULT 14  
US-10-073-865-88  
; Sequence 88, Application US/10073865  
; Publication No. US20030044904A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PTZ011  
; CURRENT APPLICATION NUMBER: US/09/764,898  
; CURRENT FILING DATE: 2001-01-17  
; PRIOR APPLICATION DATA REMOVED - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 311  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 262  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: SITE  
; NAME/KEY: SITE  
; LOCATION: (97)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Qy 14 WISAC----AMLLCHGSLOFTFOOFLHREGGTCEVIAHRCNCNKRTEERSQTVKCSC 69  
; LENGTH: 122  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Db 18 WQCTCDEGGGLFCDDLNIC--THSPCKNGATC-----SNSGQRSYT--CTC 62  
; LENGTH: 122

Qy 70 LPGKVAGTTRNRPSCYDASIVGKWCMEPCLGECKTLPDNSGWMCATG 121  
; LENGTH: 122  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Db 63 RPGYTG-----VDCEELSE--CDSNPCRNGGSKCQEDGYXCLCPG 103  
; LENGTH: 122

Query Match 11.5%; Score 87; DB 14; Length 188;  
Best Local Similarity 25.0%; Pred. No. 0.53;  
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

Qy 14 WISAC----AMLLCHGSLOHTFOOHLHRPEGGTCEVIAAHRCNCNKRLEERSQTVKCSC 69  
; LENGTH: 122  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Db 18 WQCTCDEGGGLFCDDLNIC--THHSPOCKNGATC-----SNSGQRSYT--CTC 62  
; LENGTH: 122

Qy 70 LPGKVAGTTRNRPSCYDASIVGKWCMEPCLGECKTLPDNSGWMCATG 121  
; LENGTH: 122  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Db 63 RPGYTG-----VDCEELSE--CDSNPCRNGGSKCQEDGYXCLCPG 103  
; LENGTH: 122

RESULT 15  
US-10-242-747-109  
; Sequence 109, Application US/10242747  
; Publication No. US2004005577A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PTZ07C1  
; CURRENT APPLICATION NUMBER: US/10/242,747  
; CURRENT FILING DATE: 2001-01-17  
; PRIOR APPLICATION DATA REMOVED - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 109  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: SITE  
; NAME/KEY: SITE  
; LOCATION: (97)

CURRENT FILING DATE: 2002-09-13  
 PRIOR APPLICATION NUMBER: 09/764,881  
 PRIOR FILING DATE: 2001-01-17  
 PRIOR APPLICATION NUMBER: 60/179,065  
 PRIOR FILING DATE: 2000-01-31  
 PRIOR APPLICATION NUMBER: 60/180,628  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: 60/214,886  
 PRIOR FILING DATE: 2000-06-28  
 PRIOR APPLICATION NUMBER: 60/217,487  
 PRIOR FILING DATE: 2000-07-11  
 PRIOR APPLICATION NUMBER: 60/225,758  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/220,963  
 PRIOR FILING DATE: 2000-07-26  
 PRIOR APPLICATION NUMBER: 60/217,496  
 PRIOR FILING DATE: 2000-07-11  
 PRIOR APPLICATION NUMBER: 60/225,447  
 PRIOR FILING DATE: 2000-08-14  
 PRIOR APPLICATION NUMBER: 60/218,290  
 PRIOR FILING DATE: 2000-07-14  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 192  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 109  
 LENGTH: 188  
 TYPE: PPT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (97)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (187)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (188)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-10-242-747-109

Query Match 11.5%; Score 87; DB 15; Length 188;  
 Best Local Similarity 25.0%; Pred. No. 0.53;  
 Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;  
 Name/Key: WISAC----AMLLCHGSLOHTFQDHLHREPEGGTCEVIAAHRCCKNKRIERSTOTVKSCC 69  
 Location: 14  
 Db 18 WQCTCDEGWGLFPQDLYNC--THASPKNGATC-----SNSGQRSYT--CTC 62

Query Match 11.5%; Score 87; DB 15; Length 188;  
 Best Local Similarity 25.0%; Pred. No. 0.53;  
 Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;  
 Name/Key: LPGKVAAGTTRNRSPCVDAASIVGKWCMEPLGECKTLPDNSGMWCATG 121  
 Location: 70  
 Db 63 RPYGTG-----VDCBLELSE--CDSNPCRNGGSKDQEDGYXCLCPG 103

Search completed: April 8, 2004, 13:23:16  
 Job time : 40 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 8, 2004, 13:14:59 ; Search time 39 Seconds  
(without alignments)

1075.998 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MAMYSAMSWLYLMSACAM.....SGWMCATGNKIKTRIHPRT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SPTREMBL\_25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rabbit:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacterioplasm:\*

17: sp\_archeap:\*

17 241 32.0 165 4 095902 homo sapien  
18 221.5 29.4 169 4 Q725A6  
19 106 14.1 3871 5 Q26911 caenorhabditis  
20 95.5 12.7 2516 11 Q7TQ52 mus musculus  
21 95.5 12.7 2526 11 Q7TQ51 mus musculus  
22 95.5 12.7 2531 11 Q8k428 mus musculus  
23 95.5 12.7 2531 11 Q7TQ50 mus musculus  
24 94.5 12.5 2352 5 061240 halocynthia  
25 93 12.3 2531 5 016004 lytechinus  
26 92.5 12.3 2146 5 Q9Vc97 drosophila  
27 89.5 11.9 2428 5 Q8tXK  
28 88.5 11.7 1212 13 Q13149 gallus gallus  
29 88.5 11.7 2447 13 Q13149 fugu rubripinnis  
30 87 11.5 597 11 Q35727 mus musculus  
31 85 11.3 669 4 075441 homo sapien  
32 85 11.3 1511 4 075412 homo sapien  
33 85 11.3 1587 4 000508 homo sapien  
34 84.5 11.2 2528 13 Q8axpo cynops pyrrhura  
35 84.5 11.2 2656 5 Q9GRNU3 paracentrotus  
36 84 11.1 450 11 Q8KOH9 qidkoh9  
37 84 11.1 844 11 Q7tSG9  
38 84 11.1 1405 11 Q8VHS2  
39 83 11.1 729 13 Q7T3M4  
40 83 11.0 305 6 Q9n028 macaca fasciata  
41 83 11.0 609 11 Q80YCS  
42 83 11.0 651 10 Q9fJE2  
43 83 11.0 674 5 Q8t4N9  
44 83 11.0 1065 11 Q810H2  
45 82.5 10.9 111 12 Q8JKE6 ectromelia  
Q8jke6 ectromelia

## ALIGNMENTS

RESULT 1  
Q7Z5A9

ID Q7Z5A9 PRELIMINARY; PRT; 133 AA.

AC Q7Z5A9; DT 01-OCT-2003 (TREMBLrel. 25, Created)

DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DR TAFAL1; OS Homo sapiens (Human).  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TAXID:9606; OX [1]

RP SEQUENCE FROM N.A.  
RA Tang Y.T., Bimonte P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;  
TAFA: A Novel Secreted Family with Homology to CC-chemokines.";  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY325114; AAR92406.1; -  
SEQUENCE 133 AA; 14901 MW; C5DD1C7E5F997386 CRC64;

Query Match 100.0%; Score 754; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 2.6e-80;  
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAMVSAMSWLYLWISACAMLICHGSLQHTPQQHHLRPEGGTCIVTAHRCGNKRNIE 60  
Db 1 MAMVSAMSWLYLWISACAMLICHGSLQHTPQQHHLRPEGGTCIVTAHRCGNKRNIE 60  
OY 61 RSQTVCSCSCLPGKVAGTTRNPSCYDASIVLGKWCMEPCLEGECKTLPDNSGMAT 120  
Db 61 RSQTVCSCSCLPGKVAGTTRNPSCYDASIVLGKWCMEPCLEGECKTLPDNSGMAT 120  
OY 121 GNKIKTRIHRPT 133  
Db 121 GNKIKTRIHRPT 133

%

Result No.	Query	Match	Length	DB	ID	Description
1	Q7Z5A9	133	4	Q7Z5A9		Q7Z5A9 homo sapien
2	Q7Z5A9	133	11	Q7TPG8		Q7tpg8 mus musculus
3	Q7Z5A9	133	4	Q8tCL8		Q8tcl8 homo sapien
4	Q7Z5A9	135	11	Q7TPG5		Q7tpg5 mus musculus
5	Q7Z5A9	140	6	Q9nOD3		Q9nod3 macaca fasciata
6	Q7Z5A9	140	4	Q96LR4		Q96lr4 homo sapien
7	Q7Z5A9	140	6	Q95k92		Q95k92 macaca fasciata
8	Q7Z5A9	131	4	Q8n3H0		Q8n3h0 homo sapien
9	Q7Z5A9	132	11	Q7TPG6		Q7tpg6 mus musculus
10	Q7Z5A9	131	11	Q7TPG7		Q7tpg7 mus musculus
11	Q7Z5A9	126	11	Q8BV02		Q8bv02 mus musculus
12	Q7Z5A9	133	4	Q7Z5A8		Q7z5a8 homo sapien
13	Q7Z5A9	125	4	Q7Z5A7		Q7z5a7 homo sapien
14	Q7Z5A9	125	11	Q91WE9		Q91we9 mus musculus
15	Q7Z5A9	132	4	Q8IXR8		Q8ixr8 homo sapien
16	Q7Z5A9	132	11	Q8C1V6		Q8c1v6 mus musculus

RESULT 2  
Q7TPG8



Db	78 CPGQVACTTRAPSCTEASTIVTQKWNCHMNPCLGECKVLPDYSSWCSGGNKVTK 137	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Qy	129 I 129	OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Db	138 V 138	OC Cercopithecidae; Macaca; [1]
RESULT 6		OX NCBI_TAXID:9541;
Q96LR4	PRELIMINARY;	RN SEQUENCE FROM N.A.
ID Q96LR4:	PRT; 140 AA.	RC TISSUE-Medulla oblongata;
DT 01-DEC-2001 (TREMBrel. 19, Created)		RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
DT 01-DEC-2001 (TREMBrel. 19, Last sequence update)		RA Suzuki Y., Sugano S., Hashimoto K.;
DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)		RA "Isolation of full-length cDNA clones from macaque brain cDNA libraries";
DE Hypothetical protein FLJ25161 (Similar to expressed sequence AW049604)		RU Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DE (TAFA4).		DR AB063062; BAB60784.; -.
OS Homo sapiens (Human).		KW Hypothetical protein.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;		SQ SEQUENCE 140 AA; 15654 MW; 17324670B6C97E19 CRC64;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		Query Match 57.7%; Score 435; DB 6; Length 140;
QX NCBI_TAXID:9606;		Best Local Similarity 62.0%; Pred. No. 6.1e-43;
BN		Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;
RP		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
SEQUENCE FROM N.A.		DB 18 WLFPLAYTLMVCCKLMSASSQHNRGAGHQQIKQGTCTVVAVERCCNNKRIBERSQTWKCS 77
RC TISSUE-Brain;		Query Match 57.7%; Score 435; DB 6; Length 140;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,		Best Local Similarity 62.0%; Pred. No. 6.1e-43;
RA Fujimori Y., Komiyama M., Suzuki Y., Hata H.,		Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,		DB 18 WLFPLAYTLMVCCKLMSASSQHNRGAGHQQIKQGTCTVVAVERCCNNKRIBERSQTWKCS 77
RA Nagai K., Isogai T., Sugano S.;		Query Match 57.7%; Score 435; DB 6; Length 140;
RT "NEDO human cDNA sequencing Project.";		Best Local Similarity 62.0%; Pred. No. 6.1e-43;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.		Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;
RN		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
RP		DB 18 WLFPLAYTLMVCCKLMSASSQHNRGAGHQQIKQGTCTVVAVERCCNNKRIBERSQTWKCS 77
SEQUENCE FROM N.A.		Query Match 57.7%; Score 435; DB 6; Length 140;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,		Best Local Similarity 62.0%; Pred. No. 6.1e-43;
RA Fujimori Y., Komiyama M., Suzuki Y., Hata H.,		Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,		DB 18 WLFPLAYTLMVCCKLMSASSQHNRGAGHQQIKQGTCTVVAVERCCNNKRIBERSQTWKCS 77
RA Nagai K., Isogai T., Sugano S.;		Query Match 57.7%; Score 435; DB 6; Length 140;
RT "NEDO human cDNA sequencing Project.";		Best Local Similarity 62.0%; Pred. No. 6.1e-43;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.		Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;
RN		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
RP		DB 18 WLFPLAYTLMVCCKLMSASSQHNRGAGHQQIKQGTCTVVAVERCCNNKRIBERSQTWKCS 77
SEQUENCE FROM N.A.		Query Match 57.7%; Score 435; DB 6; Length 140;
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;		Best Local Similarity 62.0%; Pred. No. 6.1e-43;
RA "TAFA: A Novel Secreted Family with Homology to CC-chemokines.";		Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;
RT Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.		DB 18 WLFPLAYTLMVCCKLMSASSQHNRGAGHQQIKQGTCTVVAVERCCNNKRIBERSQTWKCS 77
DR AKO57890; BAB71606.1; -.		Query Match 57.7%; Score 435; DB 6; Length 140;
DR EMBL; BC03166; AAH3166.1; -.		Best Local Similarity 62.0%; Pred. No. 6.1e-43;
DR EMBL; AY325117; AAP92409.1; -.		Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;
KW Hypothetical protein.		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
SEQUENCE 140 AA; 15682 MW; 173254FBBS26009 CRC64;		DB 18 WLFPLAYTLMVCCKLMSASSQHNRGAGHQQIKQGTCTVVAVERCCNNKRIBERSQTWKCS 77
Query Match 57.9%; Score 436.5; DB 4; Length 140;		Query Match 57.7%; Score 432; DB 4; Length 131;
Best Local Similarity 61.1%; Pred. No. 4e-43;		Best Local Similarity 61.2%; Pred. No. 1.3e-12;
Matches 80; Conservative 16; Mismatches 30; Indels 5; Gaps 2;		Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;
QY 4 VSAMSWYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
DB 8 VCAKSVLISHLWFLAYVLMVCCKLMSASSQHNRGAGHQQIKQGTCTVVAVERCCNNKRIBERSQTWKCS 67		DB 19 FIVTLW-----GKVVSANHHKAHHVYKTGTCEVVALHRCNNKRIBERSQTWKCS 68
QY 59 EERSQTVKCSCLPGKVAGTTRAPSCVDASIVGKWNCHMNPCLGECKTLPDNSGMWC 118		Query Match 57.7%; Score 435; DB 6; Length 140;
DB 68 EERSQTVKCSCLPGKVAGTTRAPSCVDASIVGKWNCHMNPCLGECKTLPDNSGMWC 127		Best Local Similarity 62.0%; Pred. No. 6.1e-43;
QY 119 ATGNKIXKTRI 129		Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;
DB 128 SSGNKVKTTKV 138		QY 9 WLYLWISACAMLLCHGSLOHTFQQHHLHRPFGGTCTVIAHRCNNKRIBERSQTWKCS 68
RESULT 7		DB 19 FIVTLW-----GKVVSANHHKAHHVYKTGTCEVVALHRCNNKRIBERSQTWKCS 68
Q95K92	PRELIMINARY;	AC Q95K92;
ID Q95K92:	PRT; 140 AA.	DT 01-DEC-2001 (TREMBrel. 19, Created)
DT 01-DEC-2001 (TREMBrel. 19, Last sequence update)		DT 01-OCT-2003 (TREMBrel. 25, Last annotation update)
DE Hypothetical protein.		DE Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OS		

Qy	129 I 129	Db	129 V 129	Db	69 CFPQYVGGTTRAAPSCVDASIVQRKWCMEPCLEBECKVLPDKKGWSSSGNKVKTR 128
<b>RESULT 9</b>					
Q7TPG6	PRELIMINARY;	PRT;	132 AA.	Qy	129 I 129
ID Q7TPG6				Db	129 V 129
AC Q7TPG6;					
DT 01-OCT-2003 (TREMBLrel. 25, Created)					
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)					
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)					
DE TAFAP3.					
OS Mus musculus (Mouse).					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OC NCBI_TaxID=10090;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=S7BL/6J;					
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;					
RT "TAFAP: A Novel Secreted Family with Homology to C-C-chemokines.";					
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.					
DR EMBL:AY325122; AAP92414.1; -					
SQ SEQUENCE 132 AA; 14426 MW; 2D60C5AE1BA3F00 CRC64;					
Query Match 57.0%; Score 429 5; DB 11; Length 132;					
Best Local Similarity 61.5%; Pred. No. 2.5e-42;					
Matches 83; Conservative 13; Mismatches 16; Indels 23; Gaps 4;					
Qy 5 SAMSWLXI--WISACAMLICHGLSLOHTPQQHHLHRP-----EGTCVTAAHCCN 54					
Db 9 SAGSWVIALCLAWLWTCP--ASASLQ-----PPTSAVNLVKQGTCTVTAAHCCN 55					
Qy 55 KNRLEERQTVKSCSLPGKVAGTRNPSCVDASIVQRKWCMEPCLEBECKTLPDNS 114					
Db 56 RNRIEERQTVKSCSLQGAVGTRAKPSQCDVDAISVQRKWCQMEPCLEGECKVLPDLS 115					
Qy 115 GMCATGNKIKTRI 129					
Db 116 GWSSGSGHKVKTRIKV 130					
<b>RESULT 10</b>					
Q7TPG7	PRELIMINARY;	PRT;	131 AA.	Qy	129 I 129
ID Q7TPG7;				Db	129 V 129
AC Q7TPG7;					
DT 01-OCT-2003 (TREMBLrel. 25, Created)					
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)					
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)					
DE TAFAP2.					
OS Mus musculus (Mouse).					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OC NCBI_TaxID=10090;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=S7BL/6J;					
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;					
RT "TAFAP: A Novel Secreted Family with Homology to C-C-chemokines.";					
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.					
DR EMBL:AY325121; AAP92413.1; -					
SQ SEQUENCE 131 AA; 14647 MW; FE70D61916A95F47 CRC64;					
Query Match 56.9%; Score 429; DB 11; Length 131;					
Best Local Similarity 61.2%; Pred. No. 2.8e-42;					
Matches 74; Conservative 12; Mismatches 25; Indels 10; Gaps 1;					
Qy 9 WVLWLNISACAMLICHGLSLOHTPQQHHLHRPEGGTCEVIAHCCNKRIEERSQTVKS 68					
Db 19 FIVTLW-----GXAVSSANHHKAHHVRGTCVVAHLRCNCNKRIEERSQTVKS 68					
Qy 69 CLPGKVAGTRNPSCVDASIVQRKWCMEPCLEGECKTLPDNSGMATGNKIKTR 128					
<b>RESULT 11</b>					
QBBV02	PRELIMINARY;	PRT;	126 AA.	Qy	129 I 129
ID QBBV02;				Db	129 V 129
AC QBBV02;					
DT 01-MAR-2003 (TREMBLrel. 23, Created)					
DT 01-JUN-2003 (TREMBLrel. 23, Last sequence update)					
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)					
DE CDDNA FUJ25161 FIS.					
GN CL30034118RJK.					
OS Mus musculus (Mouse).					
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
RC NCBI_TaxID=10090;					
RA [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=C57BL/6J; TISSUE=Head;					
RA MEDLINE=22-5468; PubMed=2466851;					
RA THE FANTOM Consortium,					
RA THE TAKEN Genome Exploration Research Group Phase I & II Team;					
RA "Analysis of the mouse transcriptome based on functional annotation of the FANTOM Consortium,"					
RA RT Nature 420:563-573 (2002);					
RA DR MGD; NCBI: C130034118RJK;					
SQ SEQUENCE 126 AA; 13975 MW; D2AA1F362D44ACBA CRC64;					
Query Match 55.0%; Score 415; DB 11; Length 126;					
Best Local Similarity 62.9%; Pred. No. 1.2e-40;					
Matches 73; Conservative 14; Mismatches 25; Indels 4; Gaps 2;					
Qy 9 WLYLWLNISACAMLICHGLSLOHTPQQHHLHRPEGGTCEVIAHCCNKRIEERSQTVK 66					
Db 13 WLLTYLWVCCKLMSMSQHHRGHHLRXP--GCVEVAVHRCNCNKRIEERSQTVK 70					
Qy 67 CSCLPGKVAGTRNPSCVDASIVQRKWCMEPCLEGECKTLPDNSGMATGN 122					
Db 71 CSCFPQYAGTTRAQPCVAAVTRWCMNPCLGEDCFLPPSGNCSSGN 126					
<b>RESULT 12</b>					
Q7Z5A8	PRELIMINARY;	PRT;	133 AA.	Qy	129 I 129
ID Q7Z5A8				Db	129 V 129
AC Q7Z5A8					
DT 01-OCT-2003 (TREMBLrel. 25, Created)					
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)					
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)					
DE TAFA3.					
OS Homo sapiens (Human).					
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
RC NCBI_TaxID=9606;					
RA [1]					
RP SEQUENCE FROM N.A.					
RA Tang Y.T., Emtege P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;					
RA "TAFA: A Novel Secreted Family with Homology to C-C-chemokines.";					
RA DR Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.					
RA DR EMBL:AY325116; AAP92408.1;					
RA SQ SEQUENCE 133 AA; 14776 MW; C074C703817D008D CRC64;					
Query Match 54.6%; Score 411.5; DB 4; Length 133;					
Best Local Similarity 58.5%; Pred. No. 3.2e-40;					
Matches 76; Conservative 15; Mismatches 26; Indels 13; Gaps 2;					
Qy 5 SAMSWL----YLWISACAMLICHGLSLOHTPQQHHLHRPEGGTCEVIAHCCNKRIEERSQTVKS 68					
Db 19 FIVTLW-----GXAVSSANHHKAHHVRGTCVVAHLRCNCNKRIEERSQTVKS 68					
Qy 69 CLPGKVAGTRNPSCVDASIVQRKWCMEPCLEGECKTLPDNSGMATGNKIKTR 128					

